

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: May 4, 2006, 06:06:17 ; Search time 39 Seconds

(without alignments)
19.737 Million cell updates/sec

Title: US-09-726-470A-2

Perfect score: 15

Sequence: 1 XXXRXLXF 8

Scoring table: BLOSUM62

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	15	100.0	22	2	SA7206
2	15	100.0	23	2	SA7192
3	15	100.0	39	2	B85990
4	15	100.0	42	2	T07581
5	15	100.0	57	2	S16587
6	15	100.0	60	2	D90971
7	15	100.0	60	2	D85744
8	15	100.0	63	2	AG1332
9	15	100.0	63	2	AG1703
10	15	100.0	77	2	JC5052
11	15	100.0	79	2	B83400
12	15	100.0	79	2	I38991
13	15	100.0	80	2	T39148
14	15	100.0	80	2	B43259
15	15	100.0	86	2	S35769
16	15	100.0	93	2	AC0592
17	15	100.0	97	2	B90725
18	15	100.0	97	2	C85576
19	15	100.0	97	2	F64809
20	15	100.0	98	2	A95329
21	15	100.0	100	2	E69846
22	15	100.0	101	2	S51384
23	15	100.0	101	2	AE1386
24	15	100.0	101	2	AG1761
25	15	100.0	104	2	A70665
26	15	100.0	109	2	G64609
27	15	100.0	110	2	C55228
28	15	100.0	113	2	S55533
29	15	100.0	113	2	S55530

30	15	100.0	113	2	S55528	Ig heavy chain V r
31	15	100.0	113	2	S55532	Ig heavy chain V r
32	15	100.0	113	2	S55531	Ig heavy chain V r
33	15	100.0	113	2	F72687	hypothetical prote
34	15	100.0	114	2	D71048	hypothetical prote
35	15	100.0	116	2	G82537	pterin-4-alpha-car
36	15	100.0	119	1	CUPSAM	amcyanin precuso
37	15	100.0	119	2	E72714	probable ribosomal
38	15	100.0	121	1	G1H0HE	Ig heavy chain V-I
39	15	100.0	122	2	H82231	hypothetical prote
40	15	100.0	128	2	E70547	hypothetical prote
41	15	100.0	129	2	T12924	conserved hypotet
42	15	100.0	129	2	H72627	hypothetical prote
43	15	100.0	133	2	D48776	polypeptide (E2/NS
44	15	100.0	133	2	S15825	agglutinin alpha c
45	15	100.0	133	2	B30242	stem cell protein
46	15	100.0	135	2	AF3551	hypothetical prote
47	15	100.0	138	2	T35211	probable lyase - S
48	15	100.0	138	2	A87192	conserved hypotet
49	15	100.0	139	2	G97080	hypothetical prote
50	15	100.0	140	2	H71903	hypothetical prote
51	15	100.0	141	2	S30832	hypothetical prote
52	15	100.0	146	2	I64011	hypothetical prote
53	15	100.0	147	2	S49526	protein kinase hom
54	15	100.0	148	2	A99268	conserved hypotet
55	15	100.0	149	2	D75019	hypothetical prote
56	15	100.0	149	2	A58801	hypothetical prote
57	15	100.0	155	2	C69002	mannose-specific 1
58	15	100.0	155	2	B28083	glutathione transf
59	15	100.0	155	2	B83771	hypothetical prote
60	15	100.0	157	2	H72751	hypothetical prote
61	15	100.0	157	2	E84987	16 kd heat shock p
62	15	100.0	157	2	D84063	hypothetical prote
63	15	100.0	158	2	E72702	hypothetical prote
64	15	100.0	159	2	I84615	coagulation factor
65	15	100.0	160	2	D70524	hypothetical prote
66	15	100.0	161	2	T71687	hypothetical prote
67	15	100.0	162	2	AE0446	rod shape-determin
68	15	100.0	163	2	T48289	hypothetical prote
69	15	100.0	164	2	C90442	conserved hypotet
70	15	100.0	168	2	F71503	probable glycerol-
71	15	100.0	173	2	C84368	hypothetical prote
72	15	100.0	176	2	I56314	interferon-alpha -
73	15	100.0	177	2	F83410	conserved hypotet
74	15	100.0	178	2	B84636	probable auxin-reg
75	15	100.0	180	2	T44944	hypothetical prote
76	15	100.0	181	2	S22793	partitioning prote
77	15	100.0	184	2	B86192	hypothetical prote
78	15	100.0	189	1	IVH014	interferon alpha-I
79	15	100.0	189	1	IVH04B	interferon alpha-I
80	15	100.0	189	1	IVH0A5	interferon alpha-I
81	15	100.0	189	1	IVH0A9	interferon alpha-5
82	15	100.0	189	2	I51970	interferon precurs
83	15	100.0	189	2	I52347	interferon alpha-M
84	15	100.0	189	2	T04499	auxin-induced prot
85	15	100.0	189	2	D71073	hypothetical prote
86	15	100.0	189	2	AC2854	acetyltransferase
87	15	100.0	192	2	S15661	(2'-5')oligo(A) sy
88	15	100.0	193	2	D85769	probable membrane
89	15	100.0	193	2	AC0692	probable membrane
90	15	100.0	193	2	E64919	probable membrane
91	15	100.0	193	2	AH0273	probable membrane
92	15	100.0	193	2	H90920	probable membrane
93	15	100.0	193	2	E82857	P1X protein XF003
94	15	100.0	196	2	F83312	imidazoleglycerol-
95	15	100.0	197	2	G97015	hypothetical prote
96	15	100.0	197	2	D82783	hypothetical prote
97	15	100.0	197	2	H85065	hypothetical prote
98	15	100.0	198	2	T04953	hypothetical prote
99	15	100.0	198	2	D97301	probable membrane
100	15	100.0	198	2	C33465	11c-1 protein C -

ALIGNMENTS

RESULT 1
S47206
T-cell receptor J-alpha wvVII.1 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 23-Jul-1999
C:Accession: S47206
R:Plaza, A.; Kono, D.H.; Theofilopoulos, A.N.
submitted to the EMBL Data Library, February 1993
A:Reference number: S40133
A:Accession: S47206
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-22 <PLA>
A:Cross-references: UNIPARC:UPI0000116127; EMBL:X71036; NID:g507043; PIDD:CAA50353.1; PI
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: T-cell receptor

Query Match 100.0%; Score 15; DB 2; Length 22;
Best Local Similarity 60.0%; Pred. No. 88;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RXLXF 8
| | |
DB 8 RALTF 12

RESULT 2
S47192
T-cell receptor J-alpha wvVII.2 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 23-Jul-1999
C:Accession: S47192
R:Plaza, A.; Kono, D.H.; Theofilopoulos, A.N.
submitted to the EMBL Data Library, February 1993
A:Reference number: S40133
A:Accession: S47192
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-23 <PLA>
A:Cross-references: UNIPARC:UPI0000116136; EMBL:X71051; NID:g506974; PIDD:CAA50368.1; PI
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: T-cell receptor

Query Match 100.0%; Score 15; DB 2; Length 23;
Best Local Similarity 60.0%; Pred. No. 92;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RXLXF 8
| | |
DB 9 RALTF 13

RESULT 3
B85990
Hypothetical protein Z4614 [imported] - Escherichia coli (strain O157:H7, substrain EDL9590)
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: B85990
A:Perera, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamoudis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: AB5480; MUID:21074935; PMID:11206551
A:Accession: B85990
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-39 <STO>
A:Cross-references: UNIPROT:Q8X417; UNIPARC:UPI000000065; GB:AE005174; NID:g12517881; F
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:

A:Gene: Z4614

Query Match 100.0%; Score 15; DB 2; Length 39;
Best Local Similarity 60.0%; Pred. No. 1.5e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RXLXF 8
| | |
DB 28 RALAF 32

RESULT 4
T07581
Hypothetical protein 429 - Japanese black pine chloroplast
C:Species: chloroplast Pinus thunbergiana (Japanese black pine)
C:Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 09-Jul-2004
C:Accession: T07581
R:Wakasugi, T.; Tsudzuki, J.; Ito, S.; Nakashima, K.; Tsudzuki, T.; Sugitara, M.
Proc. Natl. Acad. Sci. U.S.A. 91, 9794-9798, 1994
A:Title: Loss of all ndh genes as determined by sequencing the entire chloroplast genome
A:Reference number: Z16030; MUID:95024047; PMID:7937893
A:Accession: T07581
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-42 <WAK>
A:Cross-references: UNIPROT:Q33005; UNIPARC:UPI00000939E4; EMBL:D17510; NID:g529643; PI
C:Genetics:
A:Genome: chloroplast
C:Keywords: chloroplast

Query Match 100.0%; Score 15; DB 2; Length 42;
Best Local Similarity 60.0%; Pred. No. 1.6e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RXLXF 8
| | |
DB 14 RSLSF 18

RESULT 5
S16587
Hypothetical protein 1 - lamb's-quarters
C:Species: Chenopodium album (lamb's-quarters)
C:Date: 21-Nov-1997 #sequence_revision 26-May-1995 #text_change 26-May-1995
C:Accession: S16587
R:Doerfel, P.; Melhe, A.; Dolferus, R.; Boerner, T.
Plant Mol. Biol. 17, 155-156, 1991
A:Title: DNA sequence of a mitochondrial plasmid from Chenopodium album.
A:Reference number: S16587; MUID:91329724; PMID:1651127
A:Accession: S16587
A:Status: preliminary; translation not shown
A:Molecule type: DNA
A:Residues: 1-57 <DOE>
A:Cross-references: UNIPARC:UPI000017AFID; EMBL:X58911

Query Match 100.0%; Score 15; DB 2; Length 57;
Best Local Similarity 60.0%; Pred. No. 2.2e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RXLXF 8
| | |
DB 26 RLTTF 30

RESULT 6
D90971
Hypothetical protein ECs2740 [imported] - Escherichia coli (strain O157:H7, substrain F
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: D90971
R:Hawashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.;
gawara, N.; Yasunaga, T.; Kihara, S.; Shiba, T.; Hattori, M.; Shingawa, H.
DNA Res. 8, 11-22, 2001

A>Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and gend
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: D90971
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-60 <HAY>
A:Cross-references: UNIPROT:O8X8T7; UNIPARC:UPI00000D08D0; GB:BA000007; PIDN:BA836163.1;
A:Experimental source: strain O157:H7, substrain R1MD 0509952
C:Genetics:
A:Gene: EC62740

Query Match 100.0%; Score 15; DB 2; Length 60;
Best Local Similarity 60.0%; Pred. No. 2.3e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
Db 29 RALAF 33

RESULT 7

unknown protein encoded within prophage CP-93R [imported] - *Escherichia coli* (strain O1
C/Species: *Escherichia coli*
C/Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C/Accession: D85744
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
Hiller, L.; Grotbeck, E.U.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A>Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: D85744
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-60 <STO>
A:Cross-references: UNIPROT:O8X8T7; UNIPARC:UPI0000D08D0; GB:AE005174; NID:G12515365; F
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: Z2370

Query Match 100.0%; Score 15; DB 2; Length 60;
Best Local Similarity 60.0%; Pred. No. 2.3e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
Db 29 RALAF 33

RESULT 8

hypothetical protein lmo2063 [imported] - *Listeria monocytogenes* (strain EGD-e)
C/Species: *Listeria monocytogenes*
C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C/Accession: AG1332
R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fehl, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Krefelt, J.; Kuhn, M.; Kunst, F.; Kurapkac, G.; Madueno, E.; Maltounam, A.; Ma
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A>Title: Comparative genomics of *Listeria* species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AG1332
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-63 <GLA>
A:Cross-references: UNIPROT:O8Y5J7; UNIPARC:UPI00000555B3; GB:NC_003210; PIDN:CMD00141.1
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: lmo2063

Query Match 100.0%; Score 15; DB 2; Length 63;

Best Local Similarity 60.0%; Pred. No. 2.4e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
Db 39 RLTTF 43

RESULT 9

hypothetical protein lln2169 [imported] - *Listeria innocua* (strain C1p11262)
C/Species: *Listeria innocua*
C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C/Accession: AG1703
R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fehl,
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Krefelt, J.; Kuhn, M.; Kunst, F.; Kurapkac, G.; Madueno, E.; Maltounam, A.;
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlan
A>Title: Comparative genomics of *Listeria* species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AG1703
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-63 <GLA>
A:Cross-references: UNIPROT:Q929V4; UNIPARC:UPI00000CC793; GB:AL592022; PIDN:CAC97399.
A:Experimental source: strain C1p11262
C:Genetics:
A:Gene: lln2169

Query Match 100.0%; Score 15; DB 2; Length 63;
Best Local Similarity 60.0%; Pred. No. 2.4e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
Db 39 RLTTF 43

RESULT 10

hypothetical 8.9k protein - *Shigella flexneri*
C/Species: *Shigella flexneri*
C/Date: 31-Jan-1997 #sequence_revision 31-Jan-1997 #text_change 09-Jul-2004
C/Accession: JC5052
R:Venkatesan, M.M.; Alexander, W.A.; Fernandez-Prada, C.
Gene 175, 23-27, 1996
A>Title: A *Shigella flexneri* invasion plasmid gene, *ipgH*, with homology to 18629 and s
A:Reference number: JC5050; MUID:97074644; PMID:8917071
A:Accession: JC5052
A:Molecule type: DNA
A:Residues: 1-77 <VEN>
A:Cross-references: UNIPROT:Q54148; UNIPARC:UPI00000B43C1; GB:U28354; NID:G1016674; PI
A>Note: in the authors' translation, residues 5-7 are shown after residue 15, residues
C:Superfamily: DNA replication protein dnaC

Query Match 100.0%; Score 15; DB 2; Length 77;
Best Local Similarity 60.0%; Pred. No. 2.9e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
Db 49 RSLSF 53

RESULT 11

hypothetical protein PA1970 [imported] - *Pseudomonas aeruginosa* (strain PAO1)
C/Species: *Pseudomonas aeruginosa*
C/Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C/Accession: B83400
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mitzoguchi, S.D.; Warrenner, P.; Hickey, M.J.;

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A>Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10964043
A:Accession: B83400
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-79 <STO>
A:Cross-references: UNIPROT:Q912D4; UNIPARC:UPI00000C5538; GB:AE004623; GB:AE004091; NIT
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA1970

Query Match 100.0%; Score 15; DB 2; Length 79;
Best Local Similarity 60.0%; Pred. No. 3e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RXLXF 8
DB 44 RSLTF 48

RESULT 12
138991
tumorigenic conversion-associated protein CATRL - human
C:Species: Homo sapiens (man)
C>Date: 23-Feb-1996 #sequence_revision 23-Feb-1996 #text_change 09-Jul-2004
C:Accession: 138991
R:Li, D.; Noyes, I.; Shuler, C.; Milo, G.E.
Proc. Natl. Acad. Sci. U.S.A. 92, 6409-6413, 1995
A>Title: Cloning and sequencing of CATRL 3, a human gene associated with tumorigenic con
A:Reference number: 138991; MUID:95327656; PMID:7604004
A:Accession: 138991
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-79 <RES>
A:Cross-references: UNIPROT:Q13166; UNIPARC:UPI00001270C8; EMBL:U25433; NID:G896044; PIT
A:Gene: GDB:CATRL
A:Cross-references: GDB:633071; OMIM:600676
A:Map position: 16p13.3-16p13.3

Query Match 100.0%; Score 15; DB 2; Length 79;
Best Local Similarity 60.0%; Pred. No. 3e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RXLXF 8
DB 42 RALTF 46

RESULT 13
T39148
hypothetical protein SPAC8C9.11 - fission yeast (*Schizosaccharomyces pombe*)
C:Species: Schizosaccharomyces pombe
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T39148
R:Oliver, K.; Harris, D.; Barrett, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, September 1997
A:Reference number: Z21748
A:Accession: T39148
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-80 <OLI>
A:Cross-references: UNIPROT:O14280; UNIPARC:UPI000006AC4F; EMBL:Z99168; PIDN:CAB16299.1;
C:Experimental source: strain 972h-; cosmid c8C9
C:Genetics:
A:Gene: SPDB:SPAC8C9.11
A:Map position: 1
A:Introns: 20/3

Query Match 100.0%; Score 15; DB 2; Length 80;

Best Local Similarity 60.0%; Pred. No. 3e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RXLXF 8
DB 72 RSLTF 76

RESULT 14
B43259
H+-transporting two-sector ATPase (EC 3.6.3.14) chain a - *Enterococcus hirae* (fragment)
C:Species: Enterococcus hirae
C>Date: 10-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C:Accession: B43259
R:Shibata, C.; Ehara, T.; Tomura, K.; Igarashi, K.; Kobayashi, H.
J. Bacteriol. 174, 6117-6124, 1992
A>Title: Gene structure of *Enterococcus hirae* (Streptococcus faecalis) F1F0-ATPase, whi
A:Reference number: A43259; MUID:93015650; PMID:1328152
A:Accession: B43259
A:Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-80 <SHF>
A:Cross-references: UNIPROT:P43454; UNIPARC:UPI000017CD83
A:Experimental source: ATCC 9790
A>Note: sequence extracted from NCBI backbone (NCBIN:115116, NCBI:P:115124)
C:Keywords: hydrolase

Query Match 100.0%; Score 15; DB 2; Length 80;
Best Local Similarity 60.0%; Pred. No. 3e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RXLXF 8
DB 4 RSLTF 8

RESULT 15
S35769
T-cell receptor alpha chain - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 23-Jul-1999
C:Accession: S35769
R:Wedderburn, L.R.
submitted to the EMBL Data Library, June 1993
A:Reference number: S35769
A:Accession: S35769
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-86 <MED>
A:Cross-references: UNIPARC:UPI0000116542; EMBL:Z22965; NID:9312153; PIDN:CAA80538.1; P
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: T-cell receptor

Query Match 100.0%; Score 15; DB 2; Length 86;
Best Local Similarity 60.0%; Pred. No. 3.3e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RXLXF 8
DB 56 RALTF 60

RESULT 16
AC0592
probable membrane protein STY0789 (imported) - *Salmonella enterica* subsp. *enterica* sero
C:Species: *Salmonella enterica* subsp. *enterica* serovar Typh
A>Note: this species has also been called *Salmonella typhi*
C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AC0592
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Main, J.; Churcher
th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001

A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.
 A:Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serov
 A:Reference number: AB0502; MUID:21534947; PMID:11677608
 A:Accession: AC0592
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-93 <PAR>
 A:Cross-references: UNIPARC:UPI00005A119; GB:AL513382; PIDN:CAD05205.1; PID:q16501975;
 C:Genetics:
 A:Gene: ST10789

Query Match 100.0%; Score 15; DB 2; Length 93;
 Best Local Similarity 60.0%; Pred. No. 3.5e+02;
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
 Db 14 RALSF 18

RESULT 17

hypothetical protein ECs0770 [imported] - *Escherichia coli* (strain O157:H7, substrain R1
 C:Species: *Escherichia coli*
 C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
 C:Accession: B90725
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
 Sasawara, N.; Yasunaga, T.; Kihara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
 DNA Res. 8, 11-22, 2001
 A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genc
 A:Reference number: A95629; MUID:21156231; PMID:11258796
 A:Accession: B90725
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-97 <HAY>
 A:Cross-references: UNIPROT:P37343; UNIPARC:UPI000013A331; GB:BA000007; PIDN:BA034193.1;
 A:Experimental source: strain O157:H7, substrain R1MD 0509952
 C:Genetics:
 A:Gene: ECs0770

Query Match 100.0%; Score 15; DB 2; Length 97;
 Best Local Similarity 60.0%; Pred. No. 3.7e+02;
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
 Db 18 RALSF 22

RESULT 18

hypothetical protein ybge [imported] - *Escherichia coli* (strain O157:H7, substrain EDL93
 C:Species: *Escherichia coli*
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
 C:Accession: C85576
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
 Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; DiMaletta, E.; Potamousis, K.; Apodaca,
 Nature 409, 529-533, 2001
 A:Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.
 A:Reference number: A85480; MUID:21074935; PMID:11206551
 A:Accession: C85576
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-97 <STO>
 A:Cross-references: UNIPROT:P37343; UNIPARC:UPI000013A331; GB:AE005174; NID:q12513668; F
 A:Experimental source: strain O157:H7, substrain EDL933
 C:Genetics:
 A:Gene: ybge

Query Match 100.0%; Score 15; DB 2; Length 97;
 Best Local Similarity 60.0%; Pred. No. 3.7e+02;
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
 Db 18 RALSF 22

RESULT 19

ybge protein - *Escherichia coli* (strain K-12)
 C:Species: *Escherichia coli*
 C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
 C:Accession: F64809
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
 A.; Rose, D.J.; Mau, B.; Shao, Y.
 Science 277, 1453-1462, 1997
 A:Title: The complete genome sequence of *Escherichia coli* K-12.
 A:Reference number: A64720; MUID:97426617; PMID:9278503
 A:Accession: F64809
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-97 <BLAT>
 A:Cross-references: UNIPROT:P37343; UNIPARC:UPI000013A331; GB:AE000177; GB:U00096; NIT
 A:Experimental source: strain K-12, substrain MG1655
 C:Genetics:
 A:Gene: ybge

Query Match 100.0%; Score 15; DB 2; Length 97;
 Best Local Similarity 60.0%; Pred. No. 3.7e+02;
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
 Db 18 RALSF 22

RESULT 20

probable fragment of transposase protein [imported] - *Sinorhizobium meliloti* (strain 1
 C:Species: *Sinorhizobium meliloti*
 C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
 C:Accession: A95329
 R:Barnett, M.D.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bc
 ; Kaiman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K
 Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
 A:Title: Nucleotide sequence and predicted functions of the entire *Sinorhizobium meli*
 A:Reference number: A95262; MUID:21396509; PMID:11481432
 A:Accession: A95329
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-98 <KUR>
 A:Cross-references: UNIPROT:Q922E9; UNIPARC:UPI00000CB110; GB:AE006469; PIDN:AAK65195
 A:Experimental source: strain 1021, megaplasmid pSymA
 R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Huble
 pella, D.; Chant, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.;
 L.; Hyman, R.W.; Jones, T.
 Science 293, 668-672, 2001
 A:Authors: Kahn, D.; Kahn, M.L.; Kaiman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaur
 hebaull, P.; Vandendol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yen,
 A:Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.
 A:Reference number: A96039; MUID:21368234; PMID:11474104
 A:Contents: annotation
 C:Genetics:
 A:Gene: SMA0997
 A:Genome: plasmid

Query Match 100.0%; Score 15; DB 2; Length 98;
 Best Local Similarity 60.0%; Pred. No. 3.7e+02;
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
 Db 81 RALTF 85

RESULT 21

E69846 hypothetical protein yjce - *Bacillus subtilis*

C:Species: *Bacillus subtilis*
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C:Accession: E69846
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capiano, V.; Carter, N.M.; Chd
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Funo, S.; Gallizzi, A.; Gall
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Larinos,
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue
Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portet
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Ser
Akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpestra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasunoto, K.; Yata, K.; Yoshida, K
A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
A:Reference number: A69580; MUID:198044033; PMID:9384377
A:Accession: E69846
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-100 <KUN>
A:Cross-references: UNIPROT:Q31627; UNIPARC:UPI0000060231; GB:Z99110; GB:AL009126; NID:G
A:Experimental source: strain 168
C:Genetics:
A:Gene: yjce
C:Superfamily: *Bacillus subtilis* hypothetical protein yjce

Query Match 100.0%; Score 15; DB 2; Length 100;
Best Local Similarity 60.0%; Pred. No. 3.8e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RXLXF 8
| | |
DB 24 RALTF 28

RESULT 22

S51384 hypothetical protein YLR366w - yeast (*Saccharomyces cerevisiae*)
N:Alternate names: hypothetical protein L8039.7
C:Species: *Saccharomyces cerevisiae*
C:Date: 23-Feb-1995 #sequence_revision 12-May-1995 #text_change 19-Apr-2002
C:Accession: S51384
R:Du, Z.
submitted to the EMBL Data Library, December 1994
A:Description: The sequence of *S. cerevisiae* cosmid 8039.
A:Reference number: S51377
A:Accession: S51384
A:Molecule type: DNA
A:Residues: 1-101 <DUZ>
A:Cross-references: UNIPARC:UPI000011E4D6; EMBL:U19103; NID:g609404; PIDN:AAB67568.1; PI
C:Genetics:
A:Gene: MIPS:YLR366w
A:Cross-references: SGD:S0004358
A:Map position: 12R
C:Superfamily: *Saccharomyces* hypothetical protein YLR366w

Query Match 100.0%; Score 15; DB 2; Length 101;
Best Local Similarity 60.0%; Pred. No. 3.8e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RXLXF 8
| | |
DB 48 RLTTF 52

RESULT 23

AE1386

transcription regulator *ArsR* family homolog lmo2493 [imported] - *Listeria monocytogenes*

C:Species: *Listeria monocytogenes*
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: AE1386
R:Glaser, F.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecke
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurgey, O.; Entian, K.D.; Fahn, H
D.; Jones, L.M.; Karet, U.
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkac, G.; Madueno, E.; Maitournam, A.; M
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland
A:Title: Comparative genomics of *Listeria* species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AE1386
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-101 <GLA>
A:Cross-references: UNIPROT:Q8Y4F1; UNIPARC:UPI00005570B; GB:NC_003210; PIDN:CAD00571.
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: lmo2493

Query Match 100.0%; Score 15; DB 2; Length 101;
Best Local Similarity 60.0%; Pred. No. 3.8e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RXLXF 8
| | |
DB 47 RLTAF 51

RESULT 24

AG1761 transcription regulator *ArsR* family homolog lln2636 [imported] - *Listeria innocua* (stra
C:Species: *Listeria innocua*
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: AG1761
R:Glaser, F.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecke
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurgey, O.; Entian, K.D.; Fahn, H
D.; Jones, L.M.; Karet, U.
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkac, G.; Madueno, E.; Maitournam, A.; M
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland
A:Title: Comparative genomics of *Listeria* species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AG1761
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-101 <GLA>
A:Cross-references: UNIPROT:Q928A0; UNIPARC:UPI00000CC92B; GB:AL592022; PIDN:CAC97863.1
A:Experimental source: strain Clp11262
C:Genetics:
A:Gene: lln2636

Query Match 100.0%; Score 15; DB 2; Length 101;
Best Local Similarity 60.0%; Pred. No. 3.8e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RXLXF 8
| | |
DB 47 RLTTF 51

RESULT 25

A70665 probable ureB protein - *Mycobacterium tuberculosis* (strain H37RV)
C:Species: *Mycobacterium tuberculosis*
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C:Accession: A70665
R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon,
.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A>Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: A70500
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-104 <CO>
A:Cross-references: UNIPROT:P50048; UNIPARC:UPI0000137D88; GB:Z83859; GB:AL123456; NID:G
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: ureb
C:Superfamily: urease, beta subunit; urease 12k chain homology
F:3-100/Domain: urease 12k chain homology <U12>
Query Match 100.0%; Score 15; DB 2; Length 104;
Best Local Similarity 60.0%; Pred. No. 3.9e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

RESULT 26
G64609
hypothetical protein HP0719 - Helicobacter pylori (strain 26695)
C:Species: Helicobacter pylori
C:Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 09-Jul-2004
C:Accession: G64609
R:Tomb, J.F.; White, O.; Kertavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKen-
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Wathey, L.
Nature 388, 539-547, 1997
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.
A>Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A:Reference number: A64520; MUID:97394467; PMID:9252185
A:Accession: G64609
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-109 <TOM>
A:Cross-references: UNIPROT:O25421; UNIPARC:UPI00000C088A; GB:AE000585; GB:AE000511; NID

Query Match 100.0%; Score 15; DB 2; Length 109;
Best Local Similarity 60.0%; Pred. No. 4.1e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RXLXF 8
| | | |
Db 24 RSLAF 28

RESULT 27
C55228
hypothetical protein - Thermus aquaticus (fragment)
C:Species: Thermus aquaticus
C:Date: 25-Aug-1995 #sequence_revision 25-Aug-1995 #text_change 12-Jul-2004
C:Accession: C55228
R:Meinel, T.; Blanquet, S.
J. Bacteriol. 176, 7387-7390, 1994
A>Title: Characterization of the Thermus thermophilus locus encoding peptide deformylase
A:Reference number: A55228; MUID:95050326; PMID:7961514
A:Accession: C55228
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-110 <MEI>
A:Cross-references: UNIPARC:UPI000013AED; GB:X79087; NID:G602912; PIDN:CAA55697.1; PID
C:Superfamily: P-loop Kinase

Query Match 100.0%; Score 15; DB 2; Length 110;
Best Local Similarity 60.0%; Pred. No. 4.1e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RXLXF 8
| | | |

Db 59 RALAF 63

RESULT 28
S55533
Ig heavy chain V region pe25 - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 21-Jan-2000
C:Accession: S55533
R:Boettger, V.; Boettger, A.; Lane, E.B.; Spruce, B.A.
J. Mol. Biol. 247, 932-946, 1995
A>Title: Comprehensive epitope analysis of monoclonal anti-proenkephalin antibodies us-
utations in the variable region genes.
A:Reference number: S55528; MUID:95239763; PMID:7536850
A:Accession: S55533
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-113 <BOE>
A:Cross-references: UNIPARC:UPI0000116205; EMBL:X82594; NID:9854314; PIDN:CAA57930.1;
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:14-97/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 15; DB 2; Length 113;
Best Local Similarity 60.0%; Pred. No. 4.2e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RXLXF 8
| | | |
Db 83 RSLTF 87

RESULT 29
S55530
Ig heavy chain V region pe17 - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 21-Jan-2000
C:Accession: S55530
R:Boettger, V.; Boettger, A.; Lane, E.B.; Spruce, B.A.
J. Mol. Biol. 247, 932-946, 1995
A>Title: Comprehensive epitope analysis of monoclonal anti-proenkephalin antibodies us-
utations in the variable region genes.
A:Reference number: S55528; MUID:95239763; PMID:7536850
A:Accession: S55530
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-113 <BOE>
A:Cross-references: UNIPARC:UPI00001161PD; EMBL:X82586; NID:9854296; PIDN:CAA57922.1;
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:14-97/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 15; DB 2; Length 113;
Best Local Similarity 60.0%; Pred. No. 4.2e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RXLXF 8
| | | |
Db 83 RSLTF 87

RESULT 30
S55528
Ig heavy chain V region (pe16/pe14) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 21-Jan-2000
C:Accession: S55528; S55529
R:Boettger, V.; Boettger, A.; Lane, E.B.; Spruce, B.A.
J. Mol. Biol. 247, 932-946, 1995
A>Title: Comprehensive epitope analysis of monoclonal anti-proenkephalin antibodies us-
utations in the variable region genes.
A:Reference number: S55528; MUID:95239763; PMID:7536850
A:Accession: S55528

A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-113 <BOB>
A:Cross-references: UNIPARC:UPI00001161FB; EMBL:X82583; NID:9854294; PIDN:CAA57921.1; PI
A:Accession: S55529
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-113 <BO2>
A:Cross-references: UNIPARC:UPI00001161FB; EMBL:X82583; NID:9854290; PIDN:CAA57919.1; PI
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotrimer; immunoglobulin
F:14-97/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 15; DB 2; Length 113;
Best Local Similarity 60.0%; Pred. No. 4.2e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 RXLXF 8
DB 83 RSLTF 87

RESULT 31
S55532
Ig heavy chain V region p619 - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 21-Jan-2000
C:Accession: S55532
R:Boettger, V.; Boettger, A.; Lane, E.B.; Spruce, B.A.
J. Mol. Biol. 247, 932-946, 1995
A:Title: Comprehensive epitope analysis of monoclonal anti-proenkephalin antibodies usin
uations in the variable region genes.
A:Reference number: S55528; MUID:955239763; PMID:7536850
A:Accession: S55532
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-113 <BOE>
A:Cross-references: UNIPARC:UPI00001161FF; EMBL:X82588; NID:9854300; PIDN:CAA57924.1; PI
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotrimer; immunoglobulin
F:14-97/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 15; DB 2; Length 113;
Best Local Similarity 60.0%; Pred. No. 4.2e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 RXLXF 8
DB 83 RSLTF 87

RESULT 32
S55531
Ig heavy chain V region p618 - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 21-Jan-2000
C:Accession: S55531
R:Boettger, V.; Boettger, A.; Lane, E.B.; Spruce, B.A.
J. Mol. Biol. 247, 932-946, 1995
A:Title: Comprehensive epitope analysis of monoclonal anti-proenkephalin antibodies usin
uations in the variable region genes.
A:Reference number: S55528; MUID:95239763; PMID:7536850
A:Accession: S55531
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-113 <BOE>
A:Cross-references: UNIPARC:UPI00001161FE; EMBL:X82587; NID:9854298; PIDN:CAA57923.1; PI
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotrimer; immunoglobulin
F:14-97/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 15; DB 2; Length 113;
Best Local Similarity 60.0%; Pred. No. 4.2e+02;

Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 RXLXF 8
DB 83 RSLTF 87

RESULT 33
F72687
hypothetical protein APE0918 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C:Accession: F72687
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Hatake, Y.; Jih-no, K.; Takai
awa, H.; Takamly, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.;
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropy
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Accession: F72687
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-113 <KAW>
A:Cross-references: UNIPROT:Q9YD5; UNIPARC:UPI000005DD3B; DDBJ:AP000060; NID:95104188;
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE0918
C:Superfamily: Aeropyrum pernix hypothetical protein APE0918

Query Match 100.0%; Score 15; DB 2; Length 113;
Best Local Similarity 60.0%; Pred. No. 4.2e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 RXLXF 8
DB 60 RSLSF 64

RESULT 34
D71048
hypothetical protein PH1676 - Pyrococcus horikoshii
C:Species: Pyrococcus horikoshii
C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 09-Jul-2004
C:Accession: D71048
R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Hatake, Y.; Hino, Y.; Yamamoto, S.; Seki
M.; Ohtoku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguch
DNA Res. 5, 55-76, 1998
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic
A:Reference number: A71000; MUID:98344137; PMID:9679194
A:Accession: D71048
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-114 <KAW>
A:Cross-references: UNIPROT:O59326; UNIPARC:UPI00000631C1; GB:AP000006; NID:93236133; P
A:Experimental source: strain OT3
A:Note: this accession replaces an interim accession for a sequence replaced by GenBank
C:Genetics:
A:Gene: PH1676

Query Match 100.0%; Score 15; DB 2; Length 114;
Best Local Similarity 60.0%; Pred. No. 4.3e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 RXLXF 8
DB 23 RSLAF 27

RESULT 35
G82537
pferin-4-alpha-cardiolamine dehydratase XF2604 [imported] - Xylella fastidiosa (strain
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: G82537

R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequences
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A62515; MUID:20365717; PMID:10910347
A:Note: For a complete list of authors see reference number A59328 below
A:Accession: G82537
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-116 <STM>
A:Cross-references: UNIPROT:Q9PAB4; UNIPARC:UPI0000131A3A; GB:AEO04067; GB:AEO03849; NID:
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Brienes, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carter, H
as-Neto, E.; Docena, C.; El-Dorriy, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franco, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Klatjima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigt
chad, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Martino, C.L.; Marques, M.V.; Martins, E
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tsuchioka, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF2604

Query Match 100.0%; Score 15; DB 2; Length 116;
Best Local Similarity 60.0%; Pred. No. 4.5e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 4 RXLXF 8
Db 57 RTLAF 61

RESULT 36

CUPSAM

amicyanin precursor - Methyllobacterium extorquens (strain AM1)
C:Species: Methyllobacterium extorquens
C>Date: 04-Dec-1986 #sequence_revision 21-Jul-1995 #text_change 05-Oct-2004
R:Accession: A56621; A00295
C:Accession: A56621; A00295
A:Title: Nucleotide sequence of the amicyanin gene from Methyllobacterium extorquens AM1.
DNA Seq. 2, 53-55, 1991
A:Reference number: A56621; MUID:92199244; PMID:1802036
A:Accession: A56621
A:Molecule type: DNA
A:Residues: 1-119 <CH1>
A:Cross-references: UNIPROT:P04172; UNIPARC:UPI000012594E; GB:M57963; NID:G150014; PIDN:
A:Note: sequence modified after extraction from NCBI Backbone
A:Note: the authors translated the codon CAC for residue 70 as Asn
A:Note: sequence extracted from NCBI backbone (NCBI:89409, NCBIP:89412)
R:Ambley, R.P.; Tobari, J.
Biochem. J. 232, 451-457, 1985
A:Title: The primary structures of Pseudomonas AM1 amicyanin and pseudocyanin. Two new s
A:Reference number: A90327; MUID:86130354; PMID:4091802
A:Accession: A00295
A:Molecule type: protein
A:Residues: 21-119 <AMB>
A:Cross-references: UNIPARC:UPI0000171E38
C:Comment: This species of Pseudomonas, isolated as an airborne contaminant, uses compo
the true pseudomonads as well as methylotrophs.
C:Superfamily: plastocyanin/Azurin
C:Keywords: copper; electron transfer; metalloprotein; periplasmic space
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-119/Product: amicyanin #status experimental <MAT>
F:67,106,109,112/Binding site: copper (His, Cys, His, Met) (type 1) #status predicted

Query Match 100.0%; Score 15; DB 1; Length 119;
Best Local Similarity 60.0%; Pred. No. 4.5e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 4 RXLXF 8
Db 2 RALAF 6

RESULT 37

E72714
Probable ribosomal protein S24 APE1132 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 24-Sep-1999
R:Accession: E72714
C:Accession: E72714
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aerol
DNA Res. 6, 83-101, 1999
A:Reference number: A72450; MUID:99310339; PMID:10382666
A:Accession: E72714
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-119 <KAW>
A:Cross-references: UNIPARC:UPI000005DE12; DDBJ:AP000060; NID:G5104188; PIDN:BAAG0117
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE1132

Query Match 100.0%; Score 15; DB 2; Length 119;
Best Local Similarity 60.0%; Pred. No. 4.5e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 4 RXLXF 8
Db 95 RALSF 99

RESULT 38

G1H0HE

Ig heavy chain V-II region (He) - human
C:Species: Homo sapiens (man)
C>Date: 07-May-1981 #sequence_revision 07-May-1981 #text_change 09-Jul-2004
R:Accession: A02093
C:Accession: A02093
A:Title: Subgroups of amino acid sequences in the variable regions of immunoglobulin h
Proc. Natl. Acad. Sci. U.S.A. 64, 997-1003, 1969
A:Reference number: A02093; MUID:70114712; PMID:5264153
A:Accession: A02093
A:Molecule type: protein
A:Residues: 1-121 <CN>
A:Cross-references: UNIPROT:P01818; UNIPARC:UPI000012CEEE
C:Comment: This gamma-1 chain was isolated from a myeloma protein.
C:Genetics:
A:Gene: GDB:IGHV@
A:Cross-references: GDB:128528; OMIM:147070
A:Map position: 14q32.33-14q32.33
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: blocked amino end; heterotetramer; immunoglobulin
F:1-100/Domain: immunoglobulin homology <IMV>
F:1/Modified site: blocked amino end (Gln) (probably pyrrolidone carboxylic acid) #sta

Query Match 100.0%; Score 15; DB 1; Length 121;
Best Local Similarity 60.0%; Pred. No. 4.5e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 4 RXLXF 8
Db 104 RTLAF 108

RESULT 39

H82231

hypothetical protein VC1187 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C:Species: Vibrio cholerae
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004

C:Accession: H82231
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Claydon, R.A.; Gwin, M.L.; Dodson, R.J.;
Charlson, D.; Ermolaeva, M.D.; Vamthekuan, U.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, F.
1, R.R.; Mekalanos, U.J.; Venter, J.C.; Fraser, C.M.
A:Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: H82231
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-122 <HE>
A:Cross-references: UNIPROT:O9KSR9; UNIPARC:UPI00000C2EDB; GB:AE004198; GB:AE003852; NID:
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC1187
A:Map position: 1

Query Match 100.0%; Score 15; DB 2; Length 122;
Best Local Similarity 60.0%; Pred. No. 4.6e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RXLXF 8
DB 66 RFLAF 70

RESULT 40
E70547
hypothetical protein RV0546c - *Mycobacterium tuberculosis* (strain H37RV)
C:Species: *Mycobacterium tuberculosis*
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C:Accession: E70547
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
; Connor, R.; Davies, R.; Devlin, K.; Feldwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Sgarbes, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: E70547
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-128 <COL>
A:Cross-references: UNIPROT:O06412; UNIPARC:UPI0000031805; GB:Z95558; GB:AL123456; NID:G
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: RV0546c

Query Match 100.0%; Score 15; DB 2; Length 128;
Best Local Similarity 60.0%; Pred. No. 4.8e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RXLXF 8
DB 17 RSLSF 21

RESULT 41
T12924
conserved hypothetical protein bnrd1/yosM - *Bacillus subtilis* phage SPBC2
C:Species: *Bacillus subtilis* phage SPBC2
C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004
C:Accession: T12924; G69926
R:Lazarevic, V.; Duesterhoeft, A.; Soldo, B.; Hilbert, H.; Maue, C.; Karamata, D.
submitted to the EMBL Data Library, August 1997
A:Description: The complete nucleotide sequence of the *Bacillus subtilis* SPB2ac2 prophage
A:Reference number: 217583
A:Accession: T12924
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-129 <LAZ>
A:Cross-references: UNIPROT:O31876; UNIPARC:UPI0000060551; EMBL:AF020713; NID:g3025478;
R:Kunert, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Allioni, G.; Azevedo, V.; Berter

C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch.
A.; Ehrlich, S.D.; Emmerson, P.T.; Ertlin, K.D.; Errington, J.; Fabre, C.; Ferrati, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallizzi, A.; Galie
tech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Lallio, M.F.
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kunita, K.; Lapidus, A.; Lardinois
A:Authors: Lander, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Maue, S.; Muee
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelli
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon
A:Authors: Schleich, S.; Schroeter, R.; Scofield, F.; Sekiguchi, J.; Sekowska, A.; Sero
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosiato, V.; Uchiyama
T.; Withers, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida,
A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: G69926
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-129 <KUN>
A:Cross-references: UNIPARC:UPI0000060551; GB:Z99114; GB:AL009126; NID:g2634230; PIDN:C
A:Experimental source: strain 168
C:Genetics: <LA>
A:Gene: bnrd1
C:Genetics: <KU>
A:Gene: yosM
C:Superfamily: *Bacillus subtilis* conserved hypothetical protein yosM

Query Match 100.0%; Score 15; DB 2; Length 129;
Best Local Similarity 60.0%; Pred. No. 4.8e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RXLXF 8
DB 53 RFLSF 57

RESULT 42
H72627
hypothetical protein APE1480 - *Aeropyrum pernix* (strain KI)
C:Species: *Aeropyrum pernix*
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C:Accession: H72627
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Taka
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.;
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, *Aeropy*
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Accession: H72627
A:Status: preliminary
A:Experimental source: strain KI
C:Genetics:
A:Gene: APE1480
C:Superfamily: *Aeropyrum pernix* hypothetical protein APE1480

Query Match 100.0%; Score 15; DB 2; Length 129;
Best Local Similarity 60.0%; Pred. No. 4.8e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RXLXF 8
DB 18 RFLAF 22

RESULT 43
D48776
polyprotein (E2/NS1 region, HVR1, HVR2) - hepatitis C virus (fragment)
C:Species: hepatitis C virus
C:Date: 07-Apr-1994 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C:Accession: D48776
R:Higashi, Y.; Kakumu, S.; Yoshioke, K.; Wakita, T.; Mizokami, M.; Ohba, K.; Ito, Y.; I

Virology 197, 659-668, 1993
 A>Title: Dynamics of genome change in the E2/NS1 region of hepatitis C virus in vivo.
 A:Reference number: A48776; MUID:94069940; PMID:8245288
 A:Accession: D48776
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: nucleic acid
 A:Residues: 1-133 <HIG>
 A:Cross-references: UNIPROT:Q9PXT8; UNIPARC:UPI000000F8865
 A:Experimental source: subtype III, patient KS
 A>Note: sequence extracted from NCBI backbone (NCBIP:140217)
 C:Superfamily: hepatitis C virus genome polypeptide
 C:Keywords: polypeptide

Query Match
 Best Local Similarity 100.0%; Score 15; DB 2; Length 133;
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
 | | |
 13 RSLSF 17

RESULT 44
 S15825
 agglutinin alpha chain - Osage orange
 C:Species: MacLura pomifera (Osage orange)
 C>Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 31-Dec-2004
 C:Accession: S15825
 R:Young, N.M.; Johnston, R.A.Z.; Watson, D.C.
 FEBS Lett. 282, 382-384, 1991
 A>Title: The amino acid sequences of jacalin and the MacLura pomifera agglutinin.
 A:Reference number: S15824; MUID:91243835; PMID:2037053
 A:Accession: S15825
 A:Molecule type: Protein
 A:Residues: 1-133 <FEBS>
 A:Cross-references: UNIPROT:P18674; UNIPARC:UPI000011FE4
 C:Superfamily: Mannose-specific lectin
 C:Keywords: glycoprotein
 F:100/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match
 Best Local Similarity 100.0%; Score 15; DB 2; Length 133;
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
 | | |
 82 RSLTF 86

RESULT 45
 B30242
 stem cell protein ERA-1-399, retinoic acid-induced - mouse
 C:Species: Mus musculus (house mouse)
 C>Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 31-Dec-2004
 C:Accession: B30242
 R:Iarosa, G.J.; Gudas, L.J.
 Mol. Cell. Biol. 8, 3906-3917, 1988
 A>Title: Early retinoic acid-induced P9 teratocarcinoma stem cell gene ERA-1: alternate
 A:Reference number: A30242; MUID:89127233; PMID:2906112
 A:Accession: B30242
 A:Molecule type: mRNA
 A:Residues: 1-133 <LAR>
 A:Cross-references: UNIPROT:P09022; UNIPARC:UPI000002A6D; GB:M22115; NID:G193047; PIDN:
 C:Keywords: alternative splicing; DNA binding; homeobox; nucleus; transcription regulation

Query Match
 Best Local Similarity 100.0%; Score 15; DB 2; Length 133;
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
 | | |
 117 RSLSF 121

RESULT 46
 AF3551
 hypothetical protein BMEI10335 (imported) - Brucella melitensis (strain 16M)
 C:Species: Brucella melitensis
 C>Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
 C:Accession: AF3551
 R:DeVecchio, V.G.; Kapral, V.; Redkar, R.J.; Patra, G.; Mijer, C.; Los, T.; Ivanov
 ; Mazur, M.; Goldsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Let
 Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
 A>Title: The genome sequence of the facultative intracellular pathogen Brucella melit
 A:Reference number: AD3252; PMID:11756688
 A:Accession: AF3551
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-135 <KUR>
 A:Cross-references: UNIPROT:Q8YD42; UNIPARC:UPI000005843E; GB:AE08918; PIDN:AAL53577
 A:Experimental source: strain 16M
 C:Genetics:
 A:Gene: BMEI10335
 A:Map position: 11

Query Match
 Best Local Similarity 100.0%; Score 15; DB 2; Length 135;
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
 | | |
 16 RSLSF 20

RESULT 47
 T35211
 probable lyase - Streptomyces coelicolor
 C:Species: Streptomyces coelicolor
 C>Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004
 C:Accession: T35211
 R:Seeger, K.J.; Harris, D.; Parkhill, J.; Barrell, B.G.; Raeburn, M.A.
 submitted to the EMBL Data Library, September 1998
 A:Reference number: Z21572
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Accession: T35211
 A:Molecule type: DNA
 A:Residues: 1-138 <SEE>
 A:Cross-references: UNIPROT:O86701; UNIPARC:UPI000000DADCS; EMBL:AL031515; PIDN:CAA206
 A:Experimental source: strain A3(2)
 C:Genetics:
 A:Gene: SCOEDB:SC5C7.04
 C:Superfamily: Bacillus probable methylglyoxalase yurT

Query Match
 Best Local Similarity 100.0%; Score 15; DB 2; Length 138;
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
 | | |
 17 RALAF 21

RESULT 48
 A87192
 conserved hypothetical protein ML2261 (imported) - Mycobacterium leprae
 C:Species: Mycobacterium leprae
 C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
 C:Accession: A87192
 R:Coile, S.T.; Eigmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.;
 R.; Davies, R.M.; Devlin, K.; Duttoy, S.; Feltrivelli, T.; Fraser, A.; Hamlin, N.; Holoy
 eam, M.A.; Rutherford, K.M.
 Nature 409, 1007-1011, 2001
 A:Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.;
 A>Title: Massive gene decay in the leprosy bacillus.
 A:Reference number: A86909; MUID:21128732; PMID:11234002
 A:Accession: A87192

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-138 <STO>
A:Cross-references: UNIPROT:Q9CBB6, UNIPARC:UPI00000C6E51, GB:AL450380, NID:g13093763; F
C:Genetics:
A:Gene: ML2261

Db 3 RSLAF 7

Search completed: May 4, 2006, 06:10:23
Job time : 44 secs

Query Match 100.0%; Score 15; DB 2; Length 138;
Best Local Similarity 60.0%; Pred. No. 5.2e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RXLXF 8
| | | |
Db 17 RLSLF 21

RESULT 49

G97080
hypothetical protein CAC1466 [imported] - Clostridium acetobutylicum
C:Species: Clostridium acetobutylicum
C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
C:Accession: G97080
R:NoUling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: G97080
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-139 <KUP>
A:Cross-references: UNIPROT:Q97J20, UNIPARC:UPI00000CA1F2, GB:AE001437; PIDN:AAK79434.1;
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:
A:Gene: CAC1466

Query Match 100.0%; Score 15; DB 2; Length 139;
Best Local Similarity 60.0%; Pred. No. 5.2e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RXLXF 8
| | | |
Db 51 RSLAF 55

RESULT 50

H71903
hypothetical protein jhp0657 - Helicobacter pylori (strain J99)
C:Species: Helicobacter pylori
A:Variety: strain J99
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004
C:Accession: H71903
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;
Ives, C.; Gibson, R.; Weberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;
Nature 397, 176-180, 1999
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path
A:Reference number: A71800; MUID:99120557; PMID:9923682
A:Accession: H71903
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-140 <ARN>
A:Cross-references: UNIPROT:Q9ZLC3, UNIPARC:UPI00000D721F, GB:AE001498, GB:AE001439, NID
A:Experimental source: strain J99
C:Genetics:
A:Gene: jhp0657
C:Superfamily: Helicobacter pylori hypothetical protein jhp018

Query Match 100.0%; Score 15; DB 2; Length 140;
Best Local Similarity 60.0%; Pred. No. 5.2e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RXLXF 8
| | | |

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 4, 2006, 06:02:52 ; Search time 228 Seconds
(without alignments)
24.755 Million cell updates/sec

Title: US-09-726-470A-2
Perfect score: 15
Sequence: 1 XXXRXLXF 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues
Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : Uniprot 05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	15	100.0	04YQ95_PLABE	04YQ95 plasmidium
2	15	100.0	06KGV3_HUMAN	06KGV3 homo sapien
3	15	100.0	08JNC7_GALPH	08JNC7 macropodid
4	15	100.0	056W77_ARATH	056W77 arabidopsis
5	15	100.0	083JD9_SHIFL	083JD9 shigella fl
6	15	100.0	081L29_VIBPA	081L29 vibrio para
7	15	100.0	08X417_ECO57	08X417 escherichia
8	15	100.0	033005_PINTH	033005 pinus thunb
9	15	100.0	08WR26_ANOGA	08WR26 anopheles g
10	15	100.0	04T1B2_TETNG	04T1B2 tetradon n
11	15	100.0	05NM51_ZYMMO	05NM51 zymomonas m
12	15	100.0	09RAW5_YACTO	09RAW5 frankia sp.
13	15	100.0	054G45_DICDI	054G45 dictyosteli
14	15	100.0	05T1A2_ORYSA	05T1A2 oryza sativ
15	15	100.0	07UTY2_RHOBA	07UTY2 rhodospirill
16	15	100.0	08X8T7_ECO57	08X8T7 escherichia
17	15	100.0	071XV1_LISMF	071XV1 listeria mo
18	15	100.0	08Y5J7_LISMO	08Y5J7 listeria mo
19	15	100.0	0929V4_LISIN	0929V4 listeria in
20	15	100.0	071115_GVTN	071115 trichoplusi
21	15	100.0	04X8M0_PLACH	04X8M0 plasmidium
22	15	100.0	06KGE3_9CAUD	06KGE3 bacterioph
23	15	100.0	08SFT9_HOMGA	08SFT9 homatius gam
24	15	100.0	04KRC9_TETNG	04KRC9 tetradon n
25	15	100.0	05C2S0_SCHJA	05C2S0 schistosoma
26	15	100.0	09G1S8_9ZZZZ	09G1S8 uncultured
27	15	100.0	07NSN8_CHRVO	07NSN8 chromobacte
28	15	100.0	Y4UG_RHISN	Y4UG rhizodium s
29	15	100.0	06ZLK3_BURMA	06ZLK3 burkholderi
30	15	100.0	05KRA9_CORGL	05KRA9 corynebacte
31	15	100.0	05XYJ1_BORGA	05XYJ1 borrelia ga

32	15	100.0	0991R7_9ZZZZ	0991R7 uncultured
33	15	100.0	08MP23_DICDI	08MP23 dictyosteli
34	15	100.0	0714P1_ENTFC	0714P1 enterococcu
35	15	100.0	0918X7_ENTFA	0918X7 enterococcu
36	15	100.0	04QKW7_HAE18	04QKW7 haemophilus
37	15	100.0	054148_SHIFL	054148 shigella fl
38	15	100.0	08VSD1_SHIFL	08VSD1 shigella fl
39	15	100.0	08XP22_RALSO	08XP22 ralsconia s
40	15	100.0	098G90_RHILLO	098G90 rhizobium l
41	15	100.0	08JTV5_LSDV	08JTV5 lumpy skin
42	15	100.0	091MW1_LSDV	091MW1 lumpy skin
43	15	100.0	CATRI_HUMAN	CATRI human
44	15	100.0	0912D4_PSEAE	0912D4 pseudomonas
45	15	100.0	014280_SCHRO	014280 schizosacch
46	15	100.0	06ZH28_ORYSA	06ZH28 oryza sativ
47	15	100.0	05LFF2_BACFN	05LFF2 bacteroides
48	15	100.0	072HW8_THER2	072HW8 thermus the
49	15	100.0	06M9S7_PARUM	06M9S7 parachlamyd
50	15	100.0	04RR16_TETNG	04RR16 tetradon n
51	15	100.0	07R3K7_GIALA	07R3K7 giardia lam
52	15	100.0	089B55_BRAJA	089B55 bradyrhizob
53	15	100.0	04T498_TETNG	04T498 tetradon n
54	15	100.0	07RXB9_NEUCR	07RXB9 neurospora
55	15	100.0	053CV7_9GAMA	053CV7 macaca fusc
56	15	100.0	05YTL3_MOCFA	05YTL3 nocardia fa
57	15	100.0	06DSX8_ERWCT	06DSX8 erwina car
58	15	100.0	085UG3_SPERC	085UG3 sebastolobu
59	15	100.0	06PJL9_HUMAN	06PJL9 homo sapien
60	15	100.0	08M241_PRAANA	08M241 praeomys nat
61	15	100.0	08LM23_MASCO	08LM23 mastomys co
62	15	100.0	085345_9POYV	085345 vaccinia vi
63	15	100.0	052906_ORYSA	052906 oryza sativ
64	15	100.0	06YV6_ORYSA	06YV6 oryza sativ
65	15	100.0	05NN06_ZYMMO	05NN06 zymomonas m
66	15	100.0	064MG1_BACER	064MG1 bacteroides
67	15	100.0	08M242_MASCO	08M242 mastomys co
68	15	100.0	08M239_PRAANA	08M239 praeomys nat
69	15	100.0	056R4_9CAUD	056R4 enterobacte
70	15	100.0	056ZV0_ARATH	056ZV0 arabidopsis
71	15	100.0	04ZT23_PSEBY	04ZT23 pseudomonas
72	15	100.0	05PM22_SALPA	05PM22 salmonella
73	15	100.0	08Z0T8_SALTI	08Z0T8 salmonella
74	15	100.0	08M240_PRAANA	08M240 praeomys nat
75	15	100.0	032936_MYCLE	032936 mycobacteri
76	15	100.0	07MBN9_VIBVY	07MBN9 vibrio vuln
77	15	100.0	07MZK3_PHOHL	07MZK3 photorhabdu
78	15	100.0	09CM66_PASNU	09CM66 pasteurella
79	15	100.0	YBGE_ECOLI	YBGE escherichia
80	15	100.0	08TRU8_METAC	08TRU8 methanosaic
81	15	100.0	057RK8_SALCH	057RK8 salmonella
82	15	100.0	065UN5_MANSM	065UN5 manheimela
83	15	100.0	0922E9_RHIME	0922E9 rhizobium m
84	15	100.0	07VBD9_PRONA	07VBD9 procloporo
85	15	100.0	0981R5_RHILLO	0981R5 rhizobium l
86	15	100.0	08J2X7_PYRAB	08J2X7 pyrococcus
87	15	100.0	YUCE_BACSU	YUCE bacillus su
88	15	100.0	04SZMO_TETNG	04SZMO tetradon n
89	15	100.0	07LIF1_YEAST	07LIF1 saccharomyc
90	15	100.0	07WXG2_ALCEU	07WXG2 alcaligenes
91	15	100.0	06LP42_PHOOR	06LP42 photobacter
92	15	100.0	073JL7_TREDE	073JL7 treponema d
93	15	100.0	0928A0_LISIN	0928A0 listeria in
94	15	100.0	08Y4F1_LISMF	08Y4F1 listeria mo
95	15	100.0	071WTS_LISMF	071WTS listeria mo
96	15	100.0	RS24_AERPE	RS24 aeropyrum p
97	15	100.0	093KRC_YEREN	093KRC yersteinia en
98	15	100.0	05SJY5_THER8	05SJY5 thermus the
99	15	100.0	05P4C9_CHICK	05P4C9 gallus galli
100	15	100.0		

ALIGNMENTS

RESULT 1

Q4Y095 PLABE PRELIMINARY; PRT; 30 AA.
ID Q4Y095;
AC Q4Y095;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE Hypothetical protein (Fragment).
GN ORFNames=PI07776.00.0;
OS Plasmodium berghei;
OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
NCBI_TaxID=5821;
RN NUCLEOTIDE SEQUENCE.
RP Hall N., Kariis M., Raine J.D., Carlton J.M., Kool J.T.W.A.,
RA Beritman M., Florens L., Janssen C.S., Pain A., Christophides G.K.,
RA James K., Rutherford K., Harris B., Harris D., Churcher C.,
RA Quail M.A., Ormond D., Doggett J., Trueman H.E., Mendoza J.,
RA Bidwell S.J., Rajandream M.A., Carucci D.J., Yates J.R., Kafatos F.C.,
RA Jance C.J., Barrell B., Turner C.M.R., Waters A.P., Sinden R.S.,
RT "A comprehensive survey of the Plasmodium life cycle by genomic,
RT transcriptomic, and proteomic analyses."
RL Science 307:82-86(2005).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CA01003005; CAH99814.1; -; Genomic_DNA.
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 30 AA; 3718 MW; A608D7A02BCC490F CRC64;

Query Match 100.0%; Score 15; DB 2; Length 30;
Best Local Similarity 60.0%; Pred. No. 7.5e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RXLXF 8

DB 8 RLSLF 12

RESULT 2

Q6KGV3 HUMAN PRELIMINARY; PRT; 36 AA.
ID Q6KGV3;
AC Q6KGV3;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Cytochrome P450 variant 3A7 (Fragment).
GN Cytochrome P450 variant 3A7 (Fragment).
OS Homo sapiens (human).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Homo.
NCBI_TaxID=9606;
RN NUCLEOTIDE SEQUENCE.
RP MEDLINE=20578884; PubMed=11137287; DOI=10.1016/S0378-1119(00)00470-4;
RX Finta C., Zaphiropoulos P.G.,
RT "The human cytochrome P450 3A locus. Gene evolution by capture of
RT downstream exons."
RL Gene 260:13-23(2000).
DR EMBL; AF115322; AAG48617.1; -; Genomic_DNA.
DR EMBL; AF115321; AAG48617.1; JOINED; Genomic_DNA.
DR Ensemble; ENSG00000160870; Homo sapiens.
FT NON TER 1
SQ SEQUENCE 36 AA; 4354 MW; 32798EF0F1646523 CRC64;

Query Match 100.0%; Score 15; DB 2; Length 36;
Best Local Similarity 60.0%; Pred. No. 8.9e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RXLXF 8

DB 13 RSLAF 17

RESULT 3

Q8JNC7 9ALPH PRELIMINARY; PRT; 38 AA.
ID Q8JNC7;
AC Q8JNC7;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Protein kinase (Fragment).
GN Name=US3;
OS Macropodid herpesvirus 2 (dorcopsis wallaby herpesvirus).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
NCBI_TaxID=83440;
RN NUCLEOTIDE SEQUENCE.
RP Thompson D., Smith G.,
RA Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY048540; AAL13141.1; -; Genomic_DNA.
DR GO; GO:0016301; F:kinase activity; IEA.
KW Kinase.
FT NON TER 1
SQ SEQUENCE 38 AA; 4596 MW; C6708AA1DDCFE4AD CRC64;

Query Match 100.0%; Score 15; DB 2; Length 38;
Best Local Similarity 60.0%; Pred. No. 9.3e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RXLXF 8

DB 15 RSLTF 19

RESULT 4

Q56W77 ARATH PRELIMINARY; PRT; 39 AA.
ID Q56W77;
AC Q56W77;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Hypothetical protein.
OS Arabidopsis thaliana (mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
NCBI_TaxID=3702;
RN NUCLEOTIDE SEQUENCE.
RP Totoki Y., Seki M., Ishida J., Nakajima M., Enju A., Kamiya A.,
RA Narusaki M., Shin-I T., Nakagawa M., Sakamoto N., Oishi K., Kohara Y.,
RA Kobayashi M., Toyoda A., Sakaki Y., Sakurai T., Iida K., Akiyama K.,
RA Satou M., Toyoda T., Konagaya A., Carninci P., Kawai J.,
RA Hayashizaki Y., Shinozaki K.,
RT "Large-scale analysis of RIKEN Arabidopsis full-length (RAFL) cDNAs;"
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK222170; BAD95288.1; -; mRNA.
KW Hypothetical protein.
SQ SEQUENCE 39 AA; 4740 MW; 5B2A276901EC08BD CRC64;

Query Match 100.0%; Score 15; DB 2; Length 39;
Best Local Similarity 60.0%; Pred. No. 9.6e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RXLXF 8

DB 32 RLSLF 36

RESULT 5

Q83UD9 SHIFL

ID O83UD9 SHIFL PRELIMINARY; PRT; 39 AA.
AC O83UD9;
DT 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DE Hypothetical protein.
GN OrderedLocustNames=SF3292;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=301 / Serotype 2a;
RX MEDLINE=22272406; PubMed=1384590; DOI=10.1093/nac/gkf566;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Sun L., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157.";
RL Nucleic Acids Res. 30:4432-4441(2002).
DR EMBL; AF005674; AAN44756.1; -; Genomic DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 39 AA; 4694 MW; F56DF3E12CD5A19 CRC64;

Query Match 100.0%; Score 15; DB 2; Length 39;
Best Local Similarity 60.0%; Pred. No. 9.6e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 4 RXLXF 8
Db 28 RALAF 32

RESULT 6
ID O87L29 VIBPA PRELIMINARY; PRT; 39 AA.
AC O87L29;
DT 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DE Hypothetical protein VP2787.
GN OrderedLocustNames=VP2787;
OS Vibrio parahaemolyticus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=670;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=RIND 2210633 / Serotype O3:K6;
RX MEDLINE=22508454; PubMed=12620739; DOI=10.1016/S0140-6736(03)12659-1;
RA Maitino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
RA Iijima Y., Najima Y., Nakano M., Yamashita A., Kubota Y., Kimura S.,
RA Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
RT distinct from that of V. cholerae.";
RL Lancet 361:743-749(2003).
DR EMBL; BA000031; BAC61050.1; -; Genomic DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 39 AA; 4483 MW; CC3BC98266A617FE CRC64;

Query Match 100.0%; Score 15; DB 2; Length 39;
Best Local Similarity 60.0%; Pred. No. 9.6e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 4 RXLXF 8
Db 26 RLSLF 30

RESULT 7

O8X417 ECOS7
ID O8X417 ECOS7 PRELIMINARY; PRT; 39 AA.
AC O8X417;
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DE complete genome.
GN OrderedLocustNames=24614;
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;
RX MEDLINE=21074935; PubMed=11206551; DOI=10.1038/35054089;
RA Pena N.T., Plunkett G., Ili, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobbeck E.J., Davis N.W., Lim A., Dimantanta E.T., Potamous K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";
RL Nature 409:529-533(2001).
DR EMBL; AF005174; AAG58382.1; -; Genomic DNA.
KW Complete proteome.
SQ SEQUENCE 39 AA; 4604 MW; 9607DF8C26905A1D CRC64;

Query Match 100.0%; Score 15; DB 2; Length 39;
Best Local Similarity 60.0%; Pred. No. 9.6e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 4 RXLXF 8
Db 28 RALAF 32

RESULT 8
ID Q33005 PINTH PRELIMINARY; PRT; 42 AA.
AC Q33005;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DE OPR42g.
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE OPR42g.
OS Pinus thunbergii (Green pine) (Japanese black pine).
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferales; Coniferales; Pinaceae; Pinus; Pinus.
OX NCBI_TaxID=3350;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92212283; PubMed=1557027;
RA Tsudzuki J., Nakashima K., Tsudzuki T., Hiraoka J., Shibata M.,
RA Wakasugi T., Sugita M.;
RT "Chloroplast DNA of black pine retains a residual inverted repeat
RT lacking rRNA genes: nucleotide sequences of trnQ, trnK, psbA, trnI and
RT trnH and the absence of rps16.";
RL Mol. Gen. Genet. 232:206-214(1992).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=95094312; PubMed=8001170; DOI=10.1007/BF00333804;
RA Tsudzuki J., Ito S., Tsudzuki T., Wakasugi T., Sugita M.;
RT "A new gene encoding rRNA pro (CGG) is present in the chloroplast
RT genome of black pine: a compilation of 32 rRNA genes from black pine
RT chloroplasts.";
RL Curr. Genet. 26:153-158(1994).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=95024047; PubMed=7937893;
RA Wakasugi T., Tsudzuki J., Ito S., Nakashima K., Tsudzuki T.,
RA Sugita M.;

"Loss of all ndh genes as determined by sequencing the entire chloroplast genome of the black pine *Pinus thunbergii*.";
Proc. Natl. Acad. Sci. U.S.A. 91:9794-9798(1994).

DR EMBL: D17510; BAA04457.1; -: Genomic_DNA.
DR F017581; F017581.
DR GO: GO:0009507; Chloroplast; IEA.
KW Chloroplast.
SQ SEQUENCE 42 AA; 4945 MW; 932D81DDA0604964 CRC64;

Query Match 100.0%; Score 15; DB 2; Length 42;
Best Local Similarity 60.0%; Pred. No. 1.e+03;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 14 RLSLF 18

RESULT 9

Q8WR26 ANOGA PRELIMINARY; PRT; 46 AA.
AC Q8WR26;
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Anopheles gambiae (African malaria mosquito).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;
OC Anophelinae; Anophelinae.
OK NCBI_Taxid=7165;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22120133; PubMed=12124367;
RA Francischetti I.M., Valenzuela J.G., Pham V.M., Garfield M.K.,
Ribeiro J.M.,
RT "Toward a catalog for the transcripts and proteins (salome) from the
RT salivary gland of the malaria vector *Anopheles gambiae*.";
RL J. Exp. Biol. 205:2429-2451(2002).
DR EMBL: AF457561; AAL68791.1; -: mRNA.
KW Hypothetical protein.
SQ SEQUENCE 46 AA; 5266 MW; 9D378CAAF35E4CD2 CRC64;

Query Match 100.0%; Score 15; DB 2; Length 46;
Best Local Similarity 60.0%; Pred. No. 1.e+03;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 16 RALSF 20

RESULT 10
Q4T1B2 TETNG PRELIMINARY; PRT; 46 AA.
AC Q4T1B2;
DT 13-SEP-2005 (TReMBLrel. 31, Created)
DT 13-SEP-2005 (TReMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TReMBLrel. 31, Last annotation update)
DE Chromosome undetermined SCAF10686, whole genome shotgun sequence.
DE (Fragment).
GN ORFNames=GSTENG00008913001;
OS Tetradon nigrivittidis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OK NCBI_Taxid=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jallou O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,

Da Silva C., Salimouat M., Levy M., Boudet N., Castellano S.,
Anhouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
Biemont C., Skalli Z., Catellico L., Poulain J., De Berardinis V.,
Rauand C., Duprat S., Broctier P., Couranceau J.P., Gouzy J.,
Rauand G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
Rauand M., Wolff J.N., Guigo R., Zody M.C., Medirov J.,
Rauand V., Schachter V., Weissenbach J., Saurin W., Scarpelli C.,
Rauand P., Lander E.S., Weissenbach J., Roest Crolius H.,
RT "Genome duplication in the teleost fish *Tetradon nigrivittidis* reveals
RT the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead to the EMBL/GenBank/DBJ databases.
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL: CAE01010686; CAP93320.1; -: Genomic_DNA.
FT NON TER 46
SQ SEQUENCE 46 AA; 5314 MW; 8E6504019D2CE69 CRC64;

Query Match 100.0%; Score 15; DB 2; Length 46;
Best Local Similarity 60.0%; Pred. No. 1.e+03;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 21 RALRF 25

RESULT 11
Q5NM51 ZYMMO PRELIMINARY; PRT; 47 AA.
AC Q5NM51;
DT 01-FEB-2005 (TReMBLrel. 29, Created)
DT 01-FEB-2005 (TReMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TReMBLrel. 29, Last annotation update)
DE Hypothetical protein.
GN OrderedLocNames=ZM01585;
OS Zymomonas mobilis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;
OC Sphingomonadaceae; Zymomonas.
OK NCBI_Taxid=542;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX STRAIN=ATCC 31821 / ZM4 / CP4;
RC PubMed=15592456; DOI=10.1038/nbt1045;
RA Seo J.-S., Chong H., Park H.S., Yoon K.-O., Jung C., Kim J.J.,
Hong J.H., Kim H., Kim J.-H., Kil J.-I., Park C.J., Oh H.-M.,
Lee J.-S., Jin S.-U., Um H.-W., Lee H.-U., Oh S.-U., Kim J.Y.,
Kang H.L., Lee S.Y., Lee K.J., Kang H.S.,
RT "The genome sequence of the ethanologenic bacterium *Zymomonas mobilis*
RT ZM4.";
RL Nat. Biotechnol. 23:63-68(2005).
DR EMBL: AE008692; AA00209.1; -: Genomic_DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 47 AA; 5367 MW; B3B30EF2D3EF0395 CRC64;

Query Match 100.0%; Score 15; DB 2; Length 47;
Best Local Similarity 60.0%; Pred. No. 1.e+03;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 6 RSLAF 10

RESULT 12
Q9RAM5_9ACTO PRELIMINARY; PRT; 53 AA.
ID Q9RAM5_9ACTO PRELIMINARY;
AC Q9RAM5;

DT 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-MAY-2000 (TReMBLrel. 13, last sequence update)
 DT 01-JUN-2003 (TReMBLrel. 24, last annotation update)
 DE Putative dieneolactone hydrolase-like protein (fragment).
 OS *Frankia* sp. ACN14a/ls-r.
 OC Bacteria; Actinobacteriae; Actinobacteridae; Actinomycetales;
 OC Frankineae; Frankiaceae; Frankia.
 OX NCBI_TaxID=92643;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=ACN14a/ls-r;
 RX MEDLINE=20117996; PubMed=10652097;
 RA Pouch M.N., Cournoyer B., Baumeister W.;
 RT "Characterization of the 20S proteasome from the actinomycete
 RT *Frankia*.";
 RL Mol. Microbiol. 35:368-377(2000).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=ACN14a/ls-r;
 RA Pouch M.N., Cournoyer B., Baumeister W.;
 RL Submitted (Apr-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF142435; AF14269.1; -; Genomic_DNA.
 DR GO; GO:0016787; F:hydrolase activity; IEA.
 DR InterPro; IPR002925; Dieneolacton_hydro.
 DR Pfam; PF01738; DH; 1.
 KW Hydrolase.
 FT NON_TER
 SQ SEQUENCE 53 AA; 5748 MW; CAF445055B8B1B87 CRC64;
 Query Match 100.0%; Score 15; DB 2; Length 53;
 Best Local Similarity 60.0%; Pred. No. 1.3e+03;
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 RXLXF 8
 DB 41 RALAF 45

RESULT 13
 ID 054G45_DICDI PRELIMINARY; PRT; 57 AA.
 AC 054G45;
 DT 13-SEP-2005 (TReMBLrel. 31, Created)
 DT 13-SEP-2005 (TReMBLrel. 31, last sequence update)
 DT 13-SEP-2005 (TReMBLrel. 31, last annotation update)
 DE Hypothetical protein.
 GN ORFNames=DD8018876;
 OS *Dictyostelium discoideum* (Slime mold).
 OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
 OX NCBI_TaxID=44689;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=AX4;
 RA Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.-A.,
 RA Sugand B., Bertrman M., Song J., Olsen R., Szafranski K., Xu Q.,
 RA Tunggal R., Kummerfeld S., Madera M., Konfortov B.A., Rivero F.,
 RA Bankier A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P.,
 RA Plicher K., Chen G., Saunders D., Sodergren E., Davis P.,
 RA Kethornou A., Nie X., Hall N., Anjard C., Hemphill L., Bason N.,
 RA Farbrother P., Desany B., Just E., Morio T., Rost R., Churcher C.,
 RA Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,
 RA Muzny D., Moutier T., Pain A., Lu M., Harper D., Linday R.,
 RA Hauser H., James K., Quiles M., Mohan M.B., Saito T., Buchrieser C.,
 RA Wardrop A., Felder M., Thangavelu M., Johnson D., Knights A.,
 RA Louised H., Mungall K., Oliver K., Price C., Quail M.A.,
 RA Urushihara H., Hernandez J., Rabinowitsch E., Steffen D., Sanders M.,
 RA Ma J., Kohara Y., Sharp S., Simmonds M., Spiegler S., Tivey A.,
 RA Sugano S., White B., Walker D., Woodward J., Winckler T., Tanaka Y.,
 RA Shaulsky G., Schleicher M., Weinstock G., Rosenthal A., Cox E.C.,
 RA Chisholm R.L., Gibbs R., Loomis W.F., Platzner M., Kay R.A.,
 RA Williams J., Dear P.H., Noegel A.A., Barrell B., Kuspa A.;
 RT "The genome of the social amoeba *Dictyostelium discoideum*.";
 RL Nature 0:0-0(2005).

CC -! CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL; AF01000234; EAL62222.1; -; Genomic_DNA.
 KW Hypothetical protein.
 SQ SEQUENCE 57 AA; 6817 MW; C59D67947350CC13 CRC64;
 Query Match 100.0%; Score 15; DB 2; Length 57;
 Best Local Similarity 60.0%; Pred. No. 1.4e+03;
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 RXLXF 8
 DB 46 RSLRF 50

RESULT 14
 ID 05JIA2_ORYSA PRELIMINARY; PRT; 58 AA.
 AC 05JIA2;
 DT 10-MAY-2005 (TReMBLrel. 30, Created)
 DT 10-MAY-2005 (TReMBLrel. 30, last sequence update)
 DT 10-MAY-2005 (TReMBLrel. 30, last annotation update)
 DE Hypothetical protein P0480C01.16.
 GN Name=P0480C01.16;
 OS *Oryza sativa* (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoidae; Oryzeae; Oryza.
 OX NCBI_TaxID=39947;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,
 RA Wu J., Nishimura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,
 RA Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,
 RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,
 RA Hijishita S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M.,
 RA Ikono M., Itoh S., Itoh Y., Itoh Y., Iwabuchi A., Kamiya K.,
 RA Karasawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,
 RA Machita K., Maenara T., Mizuno H., Mizubayashi T., Mukai Y.,
 RA Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,
 RA Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai K., Shibata M.,
 RA Shimokawa T., Shomura A., Song J., Takazaki Y., Teraawa K., Tsuji K.,
 RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshihara R., Yukawa K.,
 RA Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,
 RA Yano M., Jiang J., Gojobori T.;
 RT "The genome sequence and structure of rice chromosome 1.";
 RL Nature 420:312-316(2002).
 DR EMBL; AP003453; BAB87755.1; -; Genomic_DNA.
 KW Hypothetical protein.
 SQ SEQUENCE 58 AA; 6494 MW; 6F1957C17710F1A3 CRC64;
 Query Match 100.0%; Score 15; DB 2; Length 58;
 Best Local Similarity 60.0%; Pred. No. 1.4e+03;
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 RXLXF 8
 DB 7 RALSF 11

RESULT 15
 ID 07UTY2_RHOBA PRELIMINARY; PRT; 60 AA.
 AC 07UTY2;
 DT 01-OCT-2003 (TReMBLrel. 25, Created)
 DT 01-OCT-2003 (TReMBLrel. 25, last sequence update)
 DT 01-OCT-2003 (TReMBLrel. 25, last annotation update)
 DE Hypothetical protein.
 GN OrderedLocustNames=RB3623;
 OS *Rhodospirillum rubrum*.
 OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
 OC Planctomycetaceae; Pirellula.

```
OX NCBI_TaxID=117;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=1;
RX MEDLINE=22735913; PubMed=12835416; DOI=10.1073/pnas.1431443100;
RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA Ludwig W., Gade D., Beck A., Borzym K., Heilmann K., Rabus R.,
RA Schlesner H., Amann R., Reinhardt R.;
RT "Complete genome sequence of the marine planktonicete Pirellula sp.
RT strain 1."
RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).
DR EMBL; BX294139; CAD73302.1; -; Genomic_DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 60 AA; 6661 MW; 22B188C360C9708F CRC64;

Query Match 100.0%; Score 15; DB 2; Length 60;
Best Local Similarity 60.0%; Pred. No. 1.4e+03;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 RXLXF 8
DB 56 RSLTF 60

RESULT 16
OQ8877 EC057 PRELIMINARY; PRT; 60 AA.
ID OQ8877; Q7ACV4;
AC 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 13-SEP-2005 (TReMBLrel. 31, Last annotation update)
DE Hypothetical protein ECs2740.
GN OrderedlocusNames=ECs2740, z2370;
OS Escherichia coli O157:H7;
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
NCBI_TaxID=83334;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;
RX MEDLINE=21074935; PubMed=11206551; DOI=10.1038/35054089;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Postel G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohemorrhagic Escherichia coli O157:H7."
RL Nature 409:529-533(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=O157:H7 / Sakai / RIMD 0509952 / EHEC;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Onishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Ikida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kunara S., Shiba H., Hattori M., Shingawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12."
RL DNA Res. 8:11-22(2001).
DR EMBL; AB005174; AAG56416.1; -; Genomic_DNA.
DR EMBL; BA000007; BAB36163.1; -; Genomic_DNA.
DR PIR; D85744; D85744.
DR PIR; D90971; D90971.
SQ Complete proteome; Hypothetical protein.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 60 AA; 6544 MW; AE13AA97B7255109 CRC64;
```

```
OY 4 RXLXF 8
| | |
| | |
```

```
DB 29 RALAF 33

RESULT 17
O71XV1 LISMF
ID O71XV1; LISMF PRELIMINARY; PRT; 63 AA.
AC O71XV1;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedlocusNames=LMO2365_2094;
OS Listeria monocytogenes (serotype 4b / strain F2365).
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
NCBI_TaxID=265669;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15115801; DOI=10.1093/nar/gkh562;
RA Nelson K.E., Fouts D.E., Mongodin E.F., Ravel J., DeBoy R.T.,
RA Peterson J.F., Rasko D.A., Angioli S.V., Gill S.R., Paulsen I.T.,
RA Kolonay J.F., White O., Nelson W.C., Niemeyer W.C., Beaman M.J.,
RA Brinkac L.M., Daugherty S.C., Dodson R.J., Durkin A.S., Madupu R.,
RA Haft D.H., Selengut J., Van Aken S.E., Khouri H.W., Fedorova N.,
RA Forbester H.A., Tran B., Katharoulis S., Wondolilling L.D., Uhlrich G.A.,
RA Bayles D.O., Luchansky J.B., Fraser C.M.;
RT "Whole genome comparisons of serotype 4b and 1/2a strains of the food-
RT borne pathogen Listeria monocytogenes reveal new insights into the
RT core genome components of this species."
RL Nucleic Acids Res. 32:2386-2395(2004).
DR EMBL; AE017329; AAT04864.1; -; Genomic_DNA.
DR TIGR; AL012365_2094; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 63 AA; 7504 MW; A31E13B6C0BD3050 CRC64;

Query Match 100.0%; Score 15; DB 2; Length 63;
Best Local Similarity 60.0%; Pred. No. 1.5e+03;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 RXLXF 8
DB 39 RSLTF 43

RESULT 18
O8Y5J7 LISMO
ID O8Y5J7; LISMO PRELIMINARY; PRT; 63 AA.
AC O8Y5J7;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Lmo2063 protein.
GN OrderedlocusNames=lmo2063;
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
NCBI_TaxID=1639;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=EGD-e / Serovar 1/2a;
RX MEDLINE=21537279; PubMed=11679669; DOI=10.1126/science.1063447;
RA Glaser P., Frangul L., Buchrieser C., Kuster C., Amend A.,
RA Baquero F., Bergh P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Darvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkat G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjati H.,
RA Nordisk G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schlutener T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehlant J., Cossart P.;
RT "Comparative genomics of Listeria species."
RL Science 294:849-852(2001).
DR EMBL; AL591982; CAD00141.1; -; Genomic_DNA.
```

```
OY 4 RXLXF 8
| | |
| | |
```

DR PIR: AG1332; AG1332.
DR Listlist: LMO2063; -
KM Complete Proteome.
SQ SEQUENCE 63 AA; 7504 MW; A31E07098B5D3050 CRC64;

Query Match 100.0%; Score 15; DB 2; Length 63;
Best Local Similarity 60.0%; Pred. No. 1.5e+03;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 RXLXF 8
DB 39 RLLRF 43

RESULT 19
ID Q929V4 LISIN PRELIMINARY; PRT; 63 AA.
AC Q929V4;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Lin2169 protein.
GN OrderedLocustNames=lin2169;
OS Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OC Listeria innocua.
OX NCBI_TaxID=1642;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CLIP 11262 / Setovar 6a;
RX MEDLINE=21537279; PubMed=11679669; DOI=10.1126/science.1063447;
RA Glaeser P., Frangoul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Checuani F., Couve E., de Daruvar A., Denoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Emilian K.-D., Fajhi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurupkat G.,
RA Madueno E., Maizouran A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordstak G., Novella S., de Padlos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schueter T., Simoes N., Thieret A.,
RA "Vaquez-Boland J.-A., Voss H., Wehlund J., Cossart P.;
RT "Comparative genomics of Listeria species";
RL Science 294:849-852(2001).
DR EMBL: AL596171; CAC97399.1; -; Genomic_DNA.
DR PIR: AG1703; AG1703.
DR Listlist: LIN2169; -
KM Complete proteome.
SQ SEQUENCE 63 AA; 7418 MW; B7A15C4291E86115 CRC64;

Query Match 100.0%; Score 15; DB 2; Length 63;
Best Local Similarity 60.0%; Pred. No. 1.5e+03;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 RXLXF 8
DB 39 RLLRF 43

RESULT 20
ID 07115 GVTN PRELIMINARY; PRT; 64 AA.
AC 07115;
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE P7.41.
GN Name=ORF507;
OS Trichoplusia ni granulosis virus (TnGV) (Trichoplusia ni
OS granulovirus).
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae; Granulovirus.
OX NCBI_TaxID=10462;
RN [1]
RP NUCLEOTIDE SEQUENCE.

RX MEDLINE=98264509; PubMed=9603347;
RA Bideshi D.K., Hice R.H., Ge B., Federici B.A.;
RT "Molecular characterization and expression of the Trichoplusia ni
RT granulovirus helicase gene";
RL J. Gen. Virol. 79:1309-1319(1998).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Bideshi D.B., Hice R.H., Ge B., Federici B.A.;
RL Submitted (NOV-1997) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF032994; AAC40853.1; -; Genomic_DNA.
SQ SEQUENCE 64 AA; 7416 MW; AA541DB3DDC74ED6 CRC64;

Query Match 100.0%; Score 15; DB 2; Length 64;
Best Local Similarity 60.0%; Pred. No. 1.5e+03;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 RXLXF 8
DB 15 RALLSF 19

RESULT 21
ID Q4X8M0 PLACH PRELIMINARY; PRT; 66 AA.
AC Q4X8M0;
DT 13-SEP-2005 (TREMBLrel. 31, Created)
DT 13-SEP-2005 (TREMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TREMBLrel. 31, Last annotation update)
DE Hypothetical protein (Fragment).
GN ORFNames=PC404933.00.0;
OS Plasmodium chabaudi.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=3625;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Hall N., Karras M., Raine J.D., Carlton J.M., Kooli T.W.A.,
RA Berriman M., Florens L., Janssen C.S., Pain A., Christophides G.K.,
RA James K., Rutherford K., Harris B., Harris D., Churcher C.,
RA Quail M.A., Ormond D., Duggett J., Trueman H.B., Mendoza J.,
RA Bidwell S.L., Rajandream M.A., Carnucci D.J., Yates J.R., Katatos F.C.,
RA Janse C.J., Barrett B., Turner C.M.R., Waters A.P., Sinden R.S.;
RT "A comprehensive survey of the Plasmodium life cycle by genomic,
RT transcriptomic, and proteomic analyses";
RL Science 307:82-86(2005).
CC CAUTION: The sequence shown here is derived from an
CC EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL: CAJ0100836; CAH86756.1; -; Genomic_DNA.
KM Hypothetical protein.
FT NON_TER 1
FT TER 66
SQ SEQUENCE 66 AA; 7233 MW; 540CB6FEA7CEC7E2 CRC64;

Query Match 100.0%; Score 15; DB 2; Length 66;
Best Local Similarity 60.0%; Pred. No. 1.5e+03;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 RXLXF 8
DB 1 RALLSF 5

RESULT 22
ID Q6KGE3_9CAUD PRELIMINARY; PRT; 66 AA.
AC Q6KGE3;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Hypothetical protein.
OS Bacteriophage Felix 01.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales.
OX NCBI_TaxID=77775;

```

RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Srikananthan N., Whitchard J.M., Pierson F.W., Kapur V., Weigt L.A.;
RT "Bacteriophage Felix O1: Genetic Characterization.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF320576; AA014689.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ
SEQUENCE 66 AA; 7831 MW; COD3134561842994 CRC64;

Query Match 100.0%; Score 15; DB 2; Length 66;
Best Local Similarity 60.0%; Pred. No. 1.6e+03;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RXLXF 8
DB 3 RSLSF 7

RESULT 23
Q8SF79 HOMGA PRELIMINARY; PRT; 68 AA.
ID Q8SF79;
AC Q8SF79;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cytochrome b (Fragment).
GN Name=Cytb;
OS Homarus gammarus (European lobster) (Homarus vulgaris).
OC Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Astacidea;
OC Nephropidae; Nephropidae; Homarus.
OX NCBI_TaxID=707;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=14730430; DOI=10.1007/s10126-002-0097-7;
RA Katsares V., Apostolidis A., Triantafyllidis A., Kouvatzi A.,
RA Triantafyllidis C.;
RT "Development of mitochondrial DNA primers for use with homarid
RT lobsters.";
RL Mar. Biotechnol. 5:469-479(2003).
CC -1- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
CC cytochrome c1 and the Rieske protein (By similarity).
DR EMBL; AF94203; AAM15924.1; -; Genomic_DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro: IPR005798; Cytb b6 C.
DR Pfam: PF00032; Cytochrom B C 1.
DR PROSITE; PSS1003; CYTB_CTER; 1.
KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
KW Respiratory chain; Transmembrane; Transport.
FT NON TER 1
SQ
SEQUENCE 68 AA; 8074 MW; 99EF029E09DC1D0 CRC64;

Query Match 100.0%; Score 15; DB 2; Length 68;
Best Local Similarity 60.0%; Pred. No. 1.6e+03;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RXLXF 8
DB 3 RSLSF 7

RESULT 24
Q4RC9_TETNG PRELIMINARY; PRT; 68 AA.
ID Q4RC9_TETNG;
AC Q4RC9;

```

```

DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome 21 SCAP15029, whole genome shotgun sequence.
GN (Fragment).
DE ORPNames=GSTNG00033014001;
OS Tetradon nigroviridis (Green puffer).
OC Tetradon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetradon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jallion O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dastiva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Blemont C., Skalli Z., Catolico L., Poulain J., De Berardinis V.,
RA Cruaud C., Duprat S., Broctier P., Couanceau J.P., Gouzy J., Bosak S.,
RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Wolff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Lander V., Schachter V., Queller F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissbach J., Roest Crolius H.;
RT "Genome duplication in the teleost fish Tetradon nigroviridis reveals
RT the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope, Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC preliminary data.
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC EMBL; CAAB01015029; CAG1153.1; -; Genomic_DNA.
FT NON TER 68
SQ
SEQUENCE 68 AA; 7223 MW; 3AF0CF0917521E6 CRC64;

Query Match 100.0%; Score 15; DB 2; Length 68;
Best Local Similarity 60.0%; Pred. No. 1.6e+03;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RXLXF 8
DB 39 RFLAF 43

RESULT 25
Q5C2S0_SCHJA PRELIMINARY; PRT; 69 AA.
ID Q5C2S0_SCHJA;
AC Q5C2S0;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Hypothetical protein.
OS Schistosoma japonicum (Blood fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;
OC Schistosomatidae; Schistosomatidae; Schistosoma.
OX NCBI_TaxID=6182;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Han Z.;
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY810166; AAX26055.1; -; mRNA.
KW Hypothetical protein.
SQ
SEQUENCE 69 AA; 8015 MW; 7B2222568B755911 CRC64;

Query Match 100.0%; Score 15; DB 2; Length 69;
Best Local Similarity 60.0%; Pred. No. 1.7e+03;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

Oy      4 RXLXF 8
Db      58 RLTSF 62

RESULT 26
O991S8 92ZZZ PRELIMINARY; PRT; 69 AA.
AC O991S8;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS unclassified organism.
OC unclassified; environmental samples.
OX NCBI_TaxID=155900;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Stokes H.W., Nield B.S., Mabbutt B.C., Nevalainen H., Holmes A.J.,
RA Gillings M.R.;
RT "Novel and diverse integron-like gene cassettes are prevalent in
RT natural environments."
RL Submitted (JAN-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF349105; AK28612.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 69 AA; 7736 MW; EB73269523CB349 CRC64;

Query Match
Best Local Similarity 100.0%; Score 15; DB 2; Length 69;
Matches 3; Conservative 0; Pred. No. 1.7e+03; Mismatches 2; Indels 0; Gaps 0;

Oy      4 RXLXF 8
Db      21 RLTSF 25

RESULT 27
O7NSN8 CHRVO PRELIMINARY; PRT; 70 AA.
ID O7NSN8;
AC O7NSN8;
DT 01-MAR-2004 (TREMBLrel. 26, Created)
DT 01-MAR-2004 (TREMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=CV3382;
OS Chromobacterium violaceum.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Chromobacterium.
OX NCBI_TaxID=536;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STPAIN-ATCC 12472 / DSM 30191;
RX MEDLINE=22882880; PubMed=14500782; DOI=10.1073/pnas.1832124100;
RA Vasconcelos A.T.R., de Almeida D.F., Hungria M., Guimaraes C.T.,
RA Antonio R.V., Almeida F.C., de Almeida L.G.P., de Almeida R.,
RA Alves-Gomes J.A., Andrade E.M., Aratipe J., de Araujo M.F.F.,
RA Ascoli-Filho S., Azevedo V., Baptista A.J., Batista L.A.M.,
RA Bordignon J., Brígido M.M., Brito C.A., Brocchi M., Burity H.A.,
RA Camargo A.A., Cardoso D.D.P., Carneiro N.P., Carraro D.M.,
RA Carvalho C.M.B., Cascardo J.C.M., Cavada B.S., Chelvre L.M.O.,
RA Creczynski-Pasa T.B., Cunha-Junior N.C., Fagundes N., Faicao C.L.,
RA Fantihaeti F., Farias I.P., Felipe N.S.S., Ferrari L.P., Ferro J.A.,
RA Ferro M.I.T., Franco G.R., Freitas N.S.A., Furlan L.R.,
RA Gazzinelli R.T., Gomes E.A., Goncalves P.R., Grangeiro T.B.,
RA Gracianaglia D., Grisard E.C., Hanna E.S., Jardim S.N., Laurino J.,
RA Leoi L.C.T., Lima L.F.A., Loureiro M.F., Lyra M.C.C.P.,
RA Madeira H.M.F., Manfro G.P., Maranhao A.O., Martins W.S.,
RA di Mauro S.M.Z., de Medeiros S.R.B., Meisner R.V., Moreira M.A.M.,
RA Nascimento F.F., Nicolau M.F., Oliveira J.G., Oliveira S.C.,
RA Paixao R.F.C., Parente J.A., Pedrosa F.O., Pena S.D.J., Pereira J.O.,
RA Pereira M., Pinto L.S.R.C., Pinto L.S., Porto J.I.R., Potrich D.P.,
RA Ramalho-Neto C.E., Reis A.M.M., Rigo L.U., Rondinelli E.,

```

```

RA Santos E.B.P., Santos F.R., Schneider M.P.C., Senanez H.N.,
RA Silva A.M.R., da Silva A.L.C., Silva D.W., Silva R., Simoes I.C.,
RA Simon D., Soares C.M.A., Soares R.B.A., Souza E.M., Souza K.R.L.,
RA Souza R.C., Steffens M.B.R., Steindel M., Teixeira S.R., Ymenyi T.,
RA Vettore A., Wasson R., Zaha A., Simpson A.J.G.;
RT "The complete genome sequence of Chromobacterium violaceum reveals
RT remarkable and exploitable bacterial adaptability."
RL Proc. Natl. Acad. Sci. U.S.A. 100:11660-11665(2003).
DR EMBL; AE016922; AAO61046.1; -; Genomic_DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 70 AA; 7585 MW; 1127465F5AF93628 CRC64;

Query Match
Best Local Similarity 100.0%; Score 15; DB 2; Length 70;
Matches 3; Conservative 0; Pred. No. 1.7e+03; Mismatches 2; Indels 0; Gaps 0;

Oy      4 RXLXF 8
Db      55 RLTSF 59

RESULT 28
Y4UG RHISN STANDARD; PRT; 71 AA.
ID Y4UG RHISN;
AC P55671;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical 7.8 kDa protein y4ug.
GN ORFNames=Y4UG;
OS Rhizobium sp. (strain NGR234).
OC Plasmid sym pNGR234a.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=394;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX MEDLINE=97305956; PubMed=9163424;
RA Freiberg C.A., Fellay R., Bairoch A., Broughton W.J., Rosenthal A.,
RA Perret X.;
RT "Molecular basis of symbiosis between Rhizobium and legumes."
RL Nature 387:394-401(1997).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=96389014; PubMed=8796346;
RA Freiberg C., Perret X., Broughton W.J., Rosenthal A.;
RT "Sequencing the 500-kb GC-rich symbiotic replicon of Rhizobium sp.
RT NGR234 using dye terminators and a thermostable 'sequenase': a
RT beginning."
RL Genome Res. 6:590-600(1996).
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; Z68203; -; NOT ANNOTATED CDS; Genomic DNA.
DR EMBL; AB000099; AAB91879.1; -; Genomic DNA.
KW Hypothetical protein; Plasmid; Transmembrane.
FT TRANSMEM 12 32 Potential.
SQ SEQUENCE 71 AA; 7769 MW; 655E2FDA41049001 CRC64;

Query Match
Best Local Similarity 100.0%; Score 15; DB 1; Length 71;
Matches 3; Conservative 0; Pred. No. 1.7e+03; Mismatches 2; Indels 0; Gaps 0;

Oy      4 RXLXF 8
Db      8 RLTSF 12

RESULT 29

```

062LK3 BURMA
ID 062LK3 BURMA PRELIMINARY; PRT; 71 AA.
AC 062LK3
DT 25-OCT-2004 (TReMBLrel. 28, Created)
DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)
DE Hypothetical protein.
GN Ordered locus names=BMA0634;
OS Burkholderia mallei (Pseudomonas mallei).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia.
NCBI_TaxID=13373;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 23344;
RX PubMed=15377793; DOI=10.1073/pnas.0403306101;
RA Niemeyer W.C., Deshazer D., Kim H.S., Tetteelin H., Nelson K.E.,
RA Feldblyum T.V., Ulrich R.L., Remington C.M., Brinkac L.M.,
RA Daugherty S.C., Davidson T.D., Deboy R.T., Dimitrov G., Dodson R.J.,
RA Dargatzis A.S., Gwin M.L., Haft D.H., Knout H.M., Kolonay J.F.,
RA Madupu R., Mohammed Y., Nelson W.C., Radune D., Romero C.M.,
RA Sarría S., Selengut J., Shambhlin C., Sullivan S.A., White O., Yu Y.,
RA Zafer N., Zhou L., Fraser C.M.;
RT "Structural flexibility in the Burkholderia mallei genome."
RL Proc. Natl. Acad. Sci. U.S.A. 101:14246-14251(2004).
DR EMBL: CP000010; -;
DR TIGR: BMA0634; -;
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 71 AA; 7537 MW; D53CF3D92034AD4 CRC64;

Query Match 100.0%; Score 15; DB 2; Length 71;
Best Local Similarity 60.0%; Pred. No. 1.7e+03;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 RXLXF 8
DB 11 RALTF 15

RESULT 30
05KRA9 CORGL
ID 05KRA9 CORGL PRELIMINARY; PRT; 74 AA.
AC 05KRA9;
DT 10-MAY-2005 (TReMBLrel. 30, Created)
DT 10-MAY-2005 (TReMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TReMBLrel. 30, Last annotation update)
DE Putative ribosomal pseudouridine synthase (Fragment).
OS Corynebacterium glutamicum (Brevibacterium flamm).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
NCBI_TaxID=1718;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=R;
RA Suzuki N., Okayama S., Nonaka H., Tsuge Y., Inui M., Yukawa H.;
RT "Large-Scale Engineering of the Corynebacterium glutamicum Genome."
RL Appl. Environ. Microbiol. 71:3369-3372(2005).
DR EMBL: AB193038; BAD84160.1; -; Genomic_DNA.
DR GO: GO:0009982; F:Pseudouridine synthase activity; IEA.
DR GO: GO:0003723; F:RNA binding; IEA.
DR GO: GO:0006396; P:RNA processing; IEA.
DR InterPro: IPR006145; Pseudou synth.
DR Pfam: PF00849; Pseudou synth_2; 1.
FT NON_TER 1
KW SEQUENCE 74 AA; 8349 MW; D04A8CE5DB8E9A86 CRC64;

Query Match 100.0%; Score 15; DB 2; Length 74;
Best Local Similarity 60.0%; Pred. No. 1.8e+03;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 RXLXF 8
DB 50 RLTTF 54

RESULT 31
05XVL BORG
ID 05XVL BORG PRELIMINARY; PRT; 75 AA.
AC 05XVL
DT 25-OCT-2004 (TReMBLrel. 28, Created)
DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)
DE Hypothetical protein.
GN ORFNames=BG290;
OS Borrelia garinii PB1.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia;
OC Borrelia burgdorferi group.
NCBI_TaxID=290434;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PB1;
RA Gloeckner G., Lehmann R., Romualdi A., Pradella S.,
RA Schultze-Spechtel U., Schilhabel M., Wilske B., Subnel J., Platzner M.;
RT "Comparative analysis of the Borrelia garinii genome."
RL Nucleic Acids Res. 32:6038-6046(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PB1;
RA Gloeckner G., Schilhabel M., Lehmann R., Platzner M.;
RT Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY722938; AAU6141.1; -; Genomic_DNA.
DR GO: GO:0005324; F:ATP binding; IEA.
DR GO: GO:0003678; F:DNA helicase activity; IEA.
DR GO: GO:0006260; P:DNA replication; IEA.
DR InterPro: IPR007694; DnaB_C.
DR ProDom: PD332834; DnaB_C/1.
KW Hypothetical protein.
SQ SEQUENCE 75 AA; 8609 MW; 92784FAB78DF9451 CRC64;

Query Match 100.0%; Score 15; DB 2; Length 75;
Best Local Similarity 60.0%; Pred. No. 1.8e+03;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 RXLXF 8
DB 47 RALAF 51

RESULT 32
099IR7 92ZZZ
ID 099IR7 92ZZZ PRELIMINARY; PRT; 75 AA.
AC 099IR7;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS uncultured organism.
OC Unclassified; environmental samples.
NCBI_TaxID=155900;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Stokes H.W., Nield B.S., Mabbutt B.C., Nevalainen H., Holmes A.J.,
RA Gilling M.R.;
RT "Novel and diverse integron-like gene cassettes are prevalent in natural environments."
RL submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF349110; AAK28623.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 75 AA; 8285 MW; B0DAC6D7E3CF39C2 CRC64;

Query Match 100.0%; Score 15; DB 2; Length 75;
Best Local Similarity 60.0%; Pred. No. 1.8e+03;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 RXLXF 8

Db 27 RALSF 31

```

RESULT 33
O8MP23_DICDI
ID O8MP23_DICDI PRELIMINARY; PRT; 76 AA.
AC O8MP23; O8T172;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 10-MAY-2005 (TReMBLrel. 30, Last annotation update)
DE Hypothetical protein.
GN O8MP23; O8T172;
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
NCBI_TaxID=44689;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AX4;
RA Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,
RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
RA Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RA Submitted (MAY-2002) to the EMBL/Genbank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AX4;
RX MEDLINE=22092622; PubMed=12097910; DOI=10.1038/nature00847;
RA Gloeckner G., Eichinger L., Szafranski K., Pachebat J.A.,
RA Bankier A.T., Dear P.H., Lehmann R., Baumgart C., Parra G.,
RA April J.F., Guigo R., Kumpf K., Tunggal B., Cox E., Quail M.A.,
RA Platzer M., Rosenthal A., Noegel A.A.;
RT "Sequence and analysis of chromosome 2 of Dictyostelium discoideum.";
RL Nature 418:79-85(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AX4;
RA Baumgart C.;
RL Submitted (MAR-2003) to the EMBL/Genbank/DBJ databases.
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AX4;
RA Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.-A.,
RA Sungang R., Bertman M., Song J., Olsen R., Szafranski K., Xu Q.,
RA Tunggal B., Kummerfeld S., Madera M., Konfortov B.A., Rivero F.,
RA Bankier A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P.,
RA Plicher K., Chen G., Saunders D., Sodergren E., Davis P.,
RA Kethornou A., Nie X., Hall N., Augard C., Hemphill L., Bason N.,
RA Fathorou P., Desany B., Just E., Morio T., Rost R., Churcher C.,
RA Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,
RA Muzny D., Mourier T., Pain A., Lu M., Harper D., Lindsay R.,
RA Hauser H., James K., Quiles M., Mohan M.B., Saito T., Buchrieser C.,
RA Wardrop A., Felder M., Thangavelu M., Johnson D., Knights A.,
RA Loulsegue H., Mungall K., Oliver K., Price C., Quail M.A.,
RA Urushihara H., Hernandez J., Rabinowitsch E., Steffen D., Sanders M.,
RA Ma J., Kohara Y., Sharp S., Stamonds M., Spiegler S., Tiley A.,
RA Sugano S., White B., Walker D., Woodward J., Winkler T., Tanaka Y.,
RA Shaulsky G., Schleicher M., Weinstein G., Rosenthal A., Cox E.C.,
RA Chisholm R.L., Gibbs R., Loomis W.F., Platzer M., Kay R.R.,
RA Williams J., Dear P.H., Noegel A.A., Barrell B., Kuspa A.;
RT "The genome of the social amoeba Dictyostelium discoideum.";
RL Nature 0:0-0(2005).
DR EMBL; AC123513; AAM4374.1; -; Genomic DNA.
DR EMBL; AC117070; AAM09344.2; -; Genomic DNA.
DR EMBL; AAF10100027; EAL70167.1; -; Genomic DNA.
KW Hypothetical protein.
SQ SEQUENCE 76 AA; 8901 MW; A4A875C0EC898A20 CRC64;

```

Query Match 100.0%; Score 15; DB 2; Length 76;
 Best Local Similarity 60.0%; Pred. No. 1.8e+03;
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

RESULT 34
O714P1_ENTFC
ID O714P1_ENTFC PRELIMINARY; PRT; 76 AA.
AC O714P1;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN O714P1;
OS Enterococcus faecium (Streptococcus faecium).
OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
NCBI_TaxID=1352;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=BM4524;
RX DOI=10.1046/j.1365-2958.2003.03771.x;
RA Departieu F., Courvalin P., Msadek T.;
RT "A six amino acid deletion, partially overlapping the VanS G2 ATP-
RT binding motif, leads to constitutive glycopeptide resistance in VanB-
RT type Enterococcus faecium.";
RL Mol. Microbiol. 50:1069-1083(2003).
DR EMBL; AF550667; A012899.1; -; Genomic DNA.
KW Hypothetical protein.
SQ SEQUENCE 76 AA; 8698 MW; C628291E1C80D050 CRC64;

```

Query Match 100.0%; Score 15; DB 2; Length 76;
 Best Local Similarity 60.0%; Pred. No. 1.8e+03;
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 4 RALXF 8
 7 RALSF 11

```

RESULT 35
O9L8X7_ENTFA
ID O9L8X7_ENTFA PRELIMINARY; PRT; 76 AA.
AC O9L8X7;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN O9L8X7;
OS Enterococcus faecalis (Streptococcus faecalis).
OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
NCBI_TaxID=1351;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=BM4382;
RX MEDLINE=20307505; PubMed=10846226;
RA Gantier F., Taouit S., Glaser P., Courvalin P., Galimand M.;
RT "Characterization of transposon Tn1549, conferring VanB-type
RT resistance in Enterococcus spp.";
RL Microbiology 146:1481-1489(2000).
DR EMBL; AF192329; AAF7236.1; -; Genomic DNA.
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 76 AA; 8698 MW; C628291E1C80D050 CRC64;

```

Query Match 100.0%; Score 15; DB 2; Length 76;
 Best Local Similarity 60.0%; Pred. No. 1.8e+03;
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 4 RALXF 8
 7 RALSF 11

```

RESULT 36
O4QKW7_HAE18
ID O4QKW7_HAE18 PRELIMINARY; PRT; 76 AA.
AC O4QKW7;

```

Db 4 RALXF 8
 67 RALSF 71

```

DT 13-SEP-2005 (TREMBlrel. 31, Created)
DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)
DT 13-SEP-2005 (TREMBlrel. 31, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=NT11521;
OS Haemophilus influenzae (strain 86-028NP).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=281310;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15968074; DOI=10.1128/JB.187.13.4627-4636.2005;
RA Harrison A., Dyer D.W., Gillaspay A., Ray W.C., Mungr R., Carson M.B.,
RA Zhong H., Gibson J., Gibson M., Johnson L.S., Lewis L., Bakaletz L.O.,
RA Munson R.S., Jr.;
RT "Genomic sequence of an otitis media isolate of nontypeable
RT Haemophilus influenzae: comparative study with H. influenzae serotype
RT d, strain KW20."
RL J. Bacteriol. 187:4627-4636(2005).
DR EMBL, CP000057; AAX88330.1; -; Genomic DNA.
KM Complete proteome; Hypothetical protein.
SQ SEQUENCE 76 AA; 8359 MW; DA2EBA01A768CF88 CRC64;

Query Match 100.0%; Score 15; DB 2; Length 76;
Best Local Similarity 60.0%; Pred. No. 1.8e+03;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 RXLXF 8
DB 15 RSLAF 19

RESULT 37
O54148 SHIFL PRELIMINARY; PRT; 77 AA.
ID O54148 SHIFL PRELIMINARY; PRT; 77 AA.
AC O54148.
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE 01-FEB-2005 (TREMBlrel. 29, Last annotation update)
DE Invasion plasmid pWR100 ipgH ORPA and ipgH ORFB genes, complete cds
DE (orf, hypothetical).
DE Name=S0089;
GN Shigella flexneri.
OS Shigella flexneri.
OC Plasmid invasion plasmid pWR100, and plasmid virulence plasmid pWR501.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=M90T-W; PLASMID=invasion plasmid pWR100;
RX MEDLINE=92167809; PubMed=1791758;
RA Venkatesan M.M., Buysse J.M., Hartman A.B.;
RT "Sequence variation in two ipaH genes of Shigella flexneri 5 and
RT homology to the IrgA-like family of proteins."
RL Mol. Microbiol. 5:2435-2445(1991).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=M90T-W; PLASMID=invasion plasmid pWR100;
RX MEDLINE=97074644; PubMed=8917071; DOI=10.1016/0378-1119(96)00115-1;
RA Venkatesan M.M., Alexander W.A., Fernandez-Prada C.;
RT "A Shigella flexneri invasion plasmid gene, ipgH, with homology to
RT IS629 and sequences encoding bacterial sugar phosphate transport
RT proteins."
RL Gene 173:23-27(1996).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX PLASMID=virulence plasmid pWR501;
RX MEDLINE=21189246; PubMed=11292750;
RX DOI=10.1128/JAI.69.5.3271-3285.2001;
RA Venkatesan M.M., Goldberg M.B., Rose D.J., Grotbeck E.J., Burland V.,
RA Blattner F.R.;
RT "Complete DNA sequence and analysis of the large virulence plasmid of
RT Shigella flexneri.";
```

```

RL Infect. Immun. 69:3271-3285(2001).
DR EMBL, U28354; AAC44576.1; -; Genomic DNA.
DR EMBL, AF348706; AAK18399.1; -; Genomic DNA.
DR PIR, JCS052; JCS052.
DR GO, GO:0005524; F:ATP binding; IEA.
DR InterPro, IPR002611; ITCB_ATPbind.
DR Pfam, PF01695; ITCB_1.
KM Hypothetical protein; Plasmid.
SQ SEQUENCE 77 AA; 8875 MW; 01CIDFDA949974C3 CRC64;

Query Match 100.0%; Score 15; DB 2; Length 77;
Best Local Similarity 60.0%; Pred. No. 1.8e+03;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 RXLXF 8
DB 49 RSLSF 53

RESULT 38
Q8VSD1 SHIFL PRELIMINARY; PRT; 77 AA.
ID Q8VSD1 SHIFL PRELIMINARY; PRT; 77 AA.
AC Q8VSD1.
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE IS100 ORF2.
GN OrderedLocustNames=CP0214;
OS Shigella flexneri.
OC Plasmid pCP301.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=301;
RX MEDLINE=22272406; PubMed=12384590; DOI=10.1093/nar/gkf566;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157."
RL Nucleic Acids Res. 30:4432-4441(2002).
RX EMBL, AF386526; AAU72411.1; -; Genomic DNA.
DR GO, GO:0005524; F:ATP binding; IEA.
DR InterPro, IPR002611; ITCB_ATPbind.
DR Pfam, PF01695; ITCB_1.
KM Complete proteome; Plasmid.
SQ SEQUENCE 77 AA; 8860 MW; 01D246C5C696EBA CRC64;

Query Match 100.0%; Score 15; DB 2; Length 77;
Best Local Similarity 60.0%; Pred. No. 1.8e+03;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 RXLXF 8
DB 49 RSLSF 53

RESULT 39
Q8XP22 RALSO PRELIMINARY; PRT; 78 AA.
ID Q8XP22 RALSO PRELIMINARY; PRT; 78 AA.
AC Q8XP22.
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Hypothetical protein Rsp1494.
GN OrderedLocustNames=RSpl494; ORFNames=RS03050;
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Plasmid megaplasmid.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
```


OC Burkholderiaceae; Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=GM11000;
R MEDLINE=21681879; PubMed=11823852; DOI=10.1038/415497a;
RA Salanoubat M., Gelin S., Artiguenave F., Gonzy J., Mangenot S.,
RA Ariat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA Chaudier M., Choise N., Claudel-Renard C., Cunac S., Demange N.,
RA Gaopin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,
RA Siglier P., Thebaud P., Whalen M., Winkler P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502(2002).
DR EMBL; AF460685; CAD18645.1; -; Genomic_DNA.
KW Complete proteome; Hypothetical protein; Plasmid.
SQ SEQUENCE 78 AA; 8533 MW; 7D73171150B8B7B5 CRC64;

Query Match 100.0%; Score 15; DB 2; Length 78;
Best Local Similarity 60.0%; Pred. No. 1.9e+03;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 4 RXLXF 8
Db 12 RSLAF 16

RESULT 40
O98G90 RHIL0 PRELIMINARY; PRT; 78 AA.
ID O98G90 RHIL0 PRELIMINARY;
AC O98G90;
DT 01-OCT-2001 (TReMBLrel. 18, Created)
DT 01-OCT-2001 (TReMBLrel. 18, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Ms13438 protein.
OS OrderedlocusNames=ms13438;
GN Rhizobium loci (Mesorhizobium loci).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneo T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Ideawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpō S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loci.";
RL DNA Res. 7:331-338(2000).
DR EMBL; BA000012; BAB50326.1; -; Genomic_DNA.
KW Complete proteome.
SQ SEQUENCE 78 AA; 8168 MW; 247DD71269947291 CRC64;

Query Match 100.0%; Score 15; DB 2; Length 78;
Best Local Similarity 60.0%; Pred. No. 1.9e+03;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 4 RXLXF 8
Db 55 RALSF 59

RESULT 41
O8JTV5 LSDV PRELIMINARY; PRT; 78 AA.
ID O8JTV5 LSDV PRELIMINARY;
AC O8JTV5;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Hypothetical protein LW046.

GN Name=LW046;
OS Lumpy skin disease virus (LSDV).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Capripoxvirus.
OX NCBI_TaxID=59509;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22710395; PubMed=12827464;
RA Kara P.D., Afonso C.L., Wallace D.B., Kutish G.F., Abolnik C., Lu Z.,
RA Vreede F.T., Taljaard L.C.F., Zsak A., Viljoen G.J., Rock D.L.;
RT "Comparative sequence analysis of the South African vaccine strain and
RT two virulent field isolates of lumpy skin disease virus.";
RL Arch. Virol. 148:1335-1356(2003).
DR EMBL; AF409138; AAN02771.1; -; Genomic_DNA.
DR InterPro; IPR006803; Pox_15.
DR Pfam; PF04713; Pox_15; 1.
DR PIRSF; PIRSF003768; VAC_15L; 1.
DR ProDom; PD012198; Pox_15; 1.
KW Hypothetical protein.
SQ SEQUENCE 78 AA; 8779 MW; 4B991CF366D4D56 CRC64;

Query Match 100.0%; Score 15; DB 2; Length 78;
Best Local Similarity 60.0%; Pred. No. 1.9e+03;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 4 RXLXF 8
Db 42 RSLTF 46

RESULT 42
O91MW1 LSDV PRELIMINARY; PRT; 78 AA.
ID O91MW1 LSDV PRELIMINARY;
AC O91MW1;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-FEB-2005 (TReMBLrel. 29, Last annotation update)
DE Hypothetical protein LSDV046 (Hypothetical protein LD046).
GN Name=LSDV046; Synonym=LD046;
OS Lumpy skin disease virus (LSDV).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Capripoxvirus.
OX NCBI_TaxID=59509;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22129495; PubMed=11435593;
RX DOI=10.1128/JVI.75.15.7122-7130.2001;
RA Tulman E.R., Afonso C.L., Lu Z., Zsak L., Kutish G.F., Rock D.L.;
RT "Genome of lumpy skin disease virus.";
RL J. Virol. 75:7122-7130(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22710395; PubMed=12827464;
RA Kara P.D., Afonso C.L., Wallace D.B., Kutish G.F., Abolnik C., Lu Z.,
RA Vreede F.T., Taljaard L.C.F., Zsak A., Viljoen G.J., Rock D.L.;
RT "Comparative sequence analysis of the South African vaccine strain and
RT two virulent field isolates of lumpy skin disease virus.";
RL Arch. Virol. 148:1335-1356(2003).
DR EMBL; AF409137; AAN02614.1; -; Genomic_DNA.
DR InterPro; IPR006803; Pox_15.
DR Pfam; PF04713; Pox_15; 1.
DR PIRSF; PIRSF003768; VAC_15L; 1.
DR ProDom; PD012198; Pox_15; 1.
KW Hypothetical protein.
SQ SEQUENCE 78 AA; 8779 MW; AB8D48FF36D4D50 CRC64;

Query Match 100.0%; Score 15; DB 2; Length 78;
Best Local Similarity 60.0%; Pred. No. 1.9e+03;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 4 RXLXF 8
Db 12 RSLAF 16

```

Db          42 RSLTF 46

RESULT 43
CATRI_HUMAN STANDARD; PRT; 79 AA.
ID CATRI_HUMAN
AC Q13166;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE CATRI tumorigenic conversion 1 protein (CATRI.3).
GN Name=CATRI;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominiidae;
OC Homo.
OX NCBI_TaxId=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Carcinoma;
RX MEDLINE=95327656; PubMed=7604004;
RA Li D., Noyes I., Shuler C., Milo G.E.;
RT "Cloning and sequencing of CATRI.3, a human gene associated with
RT tumorigenic conversion";
RL Proc. Natl. Acad. Sci. U.S.A. 92:6409-6413(1995).
CC -!- DEVELOPMENTAL STAGE: Associated with tumorigenic conversion.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; U25433; -, NOT_ANNOTATED_CDS; mRNA.
DR PIR; I38991; I38991.
DR Ensembl; ENSG00000198950; Homo sapiens.
DR HGNC; HGNC:1525; CATRI.
DR MIM; 600676; -.
SQ SEQUENCE 79 AA; 9224 MW; EC3667C059114CF3 CRC64;

Query Match          100.0%; Score 15; DB 1; Length 79;
Best Local Similarity 60.0%; Pred. No. 1.9e+03;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 RXLXF 8
DB 42 RALTF 46

RESULT 44
O912DA PSEAE PRELIMINARY; PRT; 79 AA.
ID O912DA PSEAE PRELIMINARY;
AC O912DA;
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN OrderededlocusNames=PA1970;
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxId=287;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043; DOI=10.1038/35023079;
RA Stover C.K., Bism X.-O.T., Erwin A.L., Mizoguchi S.D., Watterner P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Gardner R.L., Goltry L., Tolentino E., Weschbock-Wadman S., Yuan Y.,
RA Brody L.A., Coulter S.N., Folger K.R., Kas A., Lardig K., Lim R.M.,
RA Smith K.L., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Saiter M.H. Jr., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an

```

```

RT opportunistic pathogen.";
RL Nature 406:959-964(2000).
DR EMBL; AF004623; AAG05358.1; -; Genomic_DNA.
DR PIR; B83400; B83400.
DR Complete proteome; Hypothetical protein.
SQ SEQUENCE 79 AA; 8661 MW; E8900BF84ADCB70 CRC64;

Query Match          100.0%; Score 15; DB 2; Length 79;
Best Local Similarity 60.0%; Pred. No. 1.9e+03;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 RXLXF 8
DB 44 RSLTF 48

RESULT 45
O14280 SCHPO PRELIMINARY; PRT; 80 AA.
ID O14280 SCHPO PRELIMINARY;
AC O14280;
DT 01-NOV-1998 (TRENBLrel. 08, Created)
DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE SPAC8C9.11 protein.
GN ORFNames=SPAC8C9.11;
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxId=4896;
RN [1]
RP NUCLEOTIDE SEQUENCE (LARGE SCALE GENOMIC DNA).
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360; DOI=10.1038/nature724;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M.H., Lyne R., Stewart A.,
RA Sgouros J.G., Peat N., Hayles J., Baker S.G., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D.E., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K.D., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., O'Neill S., Mungall K.L., Murphy U.D., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E.,
RA Rutherford K.M., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M.N., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J.R., Volkart G., Aert R., Robben J., Grymoprez B.,
RA Welljens I., Vanstreels E., Rieger M., Schefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Duesterhoeft A., Fritz C., Holzer E., Moestl D.,
RA Hilbert H., Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R.,
RA Pohl T.M., Eger P., Zimmermann W., Wedler H., Wambut R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Motlier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Roehet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.U., Moreno S., Armstrong J., Forsberg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrett B.G., Nure P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
DR EMBL; Z99168; CAB16299.1; -; Genomic_DNA.
DR PIR; T39148; T39148.
DR HSP; Q8BGS2; IY9J.
DR GeneDB; Spombe; SPAC8C9.11; -.
DR GO; GO:0030528; F:transcription regulator activity; IEA.
DR InterPro; IPR002634; BOLA.
DR Pfam; PF01722; BOLA; 1.
KW Complete proteome.
SQ SEQUENCE 80 AA; 9190 MW; 9031CFB15A876B7B CRC64;

Query Match          100.0%; Score 15; DB 2; Length 80;
Best Local Similarity 60.0%; Pred. No. 1.9e+03;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 4 RXLXF 8
| | |
Db 72 RSLSF 76

RESULT 46
Q6ZH28_ORYSA PRELIMINARY; PRT; 82 AA.
AC Q6ZH28;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DE Hypothetical protein Q61399_H05.14.
GN Name=OJ1399_H05.14;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP004090; BAD07633.1; -; Genomic_DNA.
DR Gramene; Q6ZH28; -;
KM Hypothetical protein.
SQ SEQUENCE 82 AA; 8492 MW; 00BDAL4341F7BF1 CRC64;

Query Match 100.0%; Score 15; DB 2; Length 82;
Best Local Similarity 60.0%; Pred. No. 1.9e+03;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RXLXF 8
| | |
Db 33 RSLSF 37

RESULT 47
Q5LFH2_BACFN PRELIMINARY; PRT; 82 AA.
AC Q5LFH2;
DT 01-FEB-2005 (TReMBLrel. 29, Created)
DT 01-FEB-2005 (TReMBLrel. 29, Last sequence update)
DE Hypothetical protein.
GN OrderedLocustNames=BFL411;
OS Bacteroides fragilis (strain ATCC 25285 / NCTC 9343).
OC Bacteria; Bacteroidetes; Bacteroidales (class); Bacteroidales;
OC Bacteroidaceae; Bacteroides.
OX NCBI_TaxID=272559;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15746427; DOI=10.1126/science.1107008;
RA Cerdano-Tarrega A.-M., Patrick S., Crossman L.C., Blakely G.,
RA Abratt V., Lennard N., Poston I., Duerden B., Harris B., Quail M.A.,
RA Barron A., Clark L., Corton C., Doggett J., Holden M.T.G., Larke N.,
RA Line A., Lord A., Nordbertczak H., Ormond D., Price C.,
RA Rabinowitsch E., Woodward J., Bartell B.G., Parkhill J.;
RT "Extensive DNA inversions in the B. fragilis genome control variable
RT gene expression.";
RL Science 307:1463-1465(2005).
DR EMBL; CR626927; CAH07122.1; -; Genomic_DNA.
KM Complete proteome; Hypothetical protein.
SQ SEQUENCE 82 AA; 9719 MW; 8919E43065D2A0C1 CRC64;

Query Match 100.0%; Score 15; DB 2; Length 82;
Best Local Similarity 60.0%; Pred. No. 1.9e+03;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RXLXF 8
| | |
Db 28 RSLSF 32

RESULT 48
Q72HW8_THET2 PRELIMINARY; PRT; 83 AA.
AC Q72HW8;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DE Hypothetical protein.
GN OrderedLocustNames=TTCl364;
OS Thermus thermophilus (strain HB27 / ATCC BAA-163 / DSM 7039).
OC Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;
OC Thermus.
OX NCBI_TaxID=262724;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15064768; DOI=10.1038/nbt956;
RA Hanne A., Brueggemann H., Raasch C., Wietzer A., Hartsch T.,
RA Liesegang H., Johann A., Lienard T., Gohl O., Martinez-Arias R.,
RA Jacobi C., Starkuviene V., Schlenker S., Dencker S., Huber R.,
RA Klenk H.-P., Kramer W., Merkl R., Gotschalk G., Fritz H.-J.;
RT "The genome sequence of the extreme thermophile Thermus
RT thermophilus.";
RL Nat. Biotechnol. 22:547-553(2004).
DR EMBL; AE017305; AA581706.1; -; Genomic_DNA.
KM Complete proteome; Hypothetical protein.
SQ SEQUENCE 83 AA; 9672 MW; 5BBA95D0121BFAF CRC64;

Query Match 100.0%; Score 15; DB 2; Length 83;
Best Local Similarity 60.0%; Pred. No. 2e+03;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RXLXF 8
| | |
Db 28 RSLSF 32

RESULT 49
Q6M9S7_PARUM PRELIMINARY; PRT; 83 AA.
AC Q6M9S7;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DE Hypothetical protein.
GN OrderedLocustNames=PC1948;
OS Parachlamydia sp. (strain UME25) (subsp. Acanthamoeba sp.).
OC Bacteria; Chlamydiae; Chlamydiales; Parachlamydiaceae; Parachlamydia.
OX NCBI_TaxID=264201;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15073324;
RA Horn M., Collingro A., Schmitz-Esser S., Beier C.L., Purkhöld U.,
RA Fartmann B., Brandt P., Wakabayashi G.J., Droge M., Frishman D.,
RA Rattei T., Mews H.-W., Wagner M.,
RT "Illuminating the evolutionary history of chlamydiae.";
RL Science 304:728-730(2004).
DR EMBL; BX908798; CAF24672.1; -; Genomic_DNA.
KM Complete proteome; Hypothetical protein.
SQ SEQUENCE 83 AA; 9855 MW; 35F7CC84585D7B99 CRC64;

Query Match 100.0%; Score 15; DB 2; Length 83;
Best Local Similarity 60.0%; Pred. No. 2e+03;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RXLXF 8
| | |
Db 57 RSLSF 61

RESULT 50
Q4RI6_TETNG

ID OARR16 TETNG PRELIMINARY; PRT; 84 AA.
AC OARR16-
DT 13-SEP-2005 (TREMBlrel. 31, Created)
DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)
DT 13-SEP-2005 (TREMBlrel. 31, Last annotation update)
DE Chromosome 16 SCAFI5002, whole genome shotgun sequence.
GN ORFNames=STENG00030155001;
NM
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jallion O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Bismont C., Skallil Z., Catolico L., Poulain J., De Bernardis V.,
RA Cruaud C., Duprat S., Brotier P., Coutanceau J.P., Gonzy J.,
RA Parra G., Lardier G., Chapelle C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff J.N., Guigo R., Zody M.C., Westrov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissbach J., Roest Croliis H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAE01015002; GAG08996.1; -; Genomic_DNA.
SQ SEQUENCE 84 AA; 9588 MW; 47E2B9B30EA67D5F CRC64;

Query Match 100.0%; Score 15; DB 2; Length 84;
Best Local Similarity 60.0%; Pred. No. 2e+03;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 RXLXF 8
| | |
Db 6 RTLSF 10

Search completed: May 4, 2006, 06:09:43
Job time : 237 secs

Chrism.B.
09/726479 Page 1
Seq. 10 2

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 4, 2006, 06:02:18 ; Search time 185 Seconds

(without alignments)
19.000 Million cell updates/sec

Title: US-09-726-470A-2

Perfect score: 15

Sequence: 1 XXXXXLKF 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :
1: Genesep21:*
2: Genesep1980s:*
3: Genesep1990s:*
4: Genesep2000s:*
5: Genesep2002s:*
6: Genesep2003as:*
7: Genesep2003bs:*
8: Genesep2004s:*
9: Genesep2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	15	100.0	5	9	ADZ71692 p21-deriv
2	15	100.0	5	9	ADZ72055 p21-deriv
3	15	100.0	10	7	ADM37109 Mutant Bc
4	15	100.0	10	7	ADM53404 Bcl-2 mut
5	15	100.0	11	7	ADZ82927 Protein k
6	15	100.0	11	9	ADZ82927 Protein k
7	15	100.0	11	9	ADZ82927 Protein k
8	15	100.0	12	9	ADZ82927 Protein k
9	15	100.0	13	7	ADZ82927 Protein k
10	15	100.0	15	2	ADZ82927 Protein k
11	15	100.0	17	6	ADZ82927 Protein k
12	15	100.0	18	6	ADZ82927 Protein k
13	15	100.0	19	8	ADZ82927 Protein k
14	15	100.0	20	5	ADZ82927 Protein k
15	15	100.0	20	6	ADZ82927 Protein k
16	15	100.0	20	9	ADZ82927 Protein k
17	15	100.0	21	2	ADZ82927 Protein k
18	15	100.0	21	8	ADZ82927 Protein k
19	15	100.0	27	3	ADZ82927 Protein k
20	15	100.0	27	5	ADZ82927 Protein k
21	15	100.0	28	3	ADZ82927 Protein k
22	15	100.0	30	2	ADZ82927 Protein k
23	15	100.0	30	7	ADZ82927 Protein k
24	15	100.0	30	8	ADZ82927 Protein k

25	15	100.0	31	2	AAW45039
26	15	100.0	31	2	AAW45039
27	15	100.0	31	2	AAW45039
28	15	100.0	31	2	AAW45039
29	15	100.0	31	2	AAW45039
30	15	100.0	31	2	AAW45039
31	15	100.0	31	2	AAW45039
32	15	100.0	31	2	AAW45039
33	15	100.0	31	2	AAW45039
34	15	100.0	31	2	AAW45039
35	15	100.0	31	2	AAW45039
36	15	100.0	31	2	AAW45039
37	15	100.0	31	2	AAW45039
38	15	100.0	31	2	AAW45039
39	15	100.0	31	2	AAW45039
40	15	100.0	31	2	AAW45039
41	15	100.0	31	2	AAW45039
42	15	100.0	31	2	AAW45039
43	15	100.0	31	2	AAW45039
44	15	100.0	31	2	AAW45039
45	15	100.0	31	2	AAW45039
46	15	100.0	31	2	AAW45039
47	15	100.0	31	2	AAW45039
48	15	100.0	31	2	AAW45039
49	15	100.0	31	2	AAW45039
50	15	100.0	31	2	AAW45039
51	15	100.0	31	2	AAW45039
52	15	100.0	31	2	AAW45039
53	15	100.0	31	2	AAW45039
54	15	100.0	31	2	AAW45039
55	15	100.0	31	2	AAW45039
56	15	100.0	31	2	AAW45039
57	15	100.0	31	2	AAW45039
58	15	100.0	31	2	AAW45039
59	15	100.0	31	2	AAW45039
60	15	100.0	31	2	AAW45039
61	15	100.0	31	2	AAW45039
62	15	100.0	31	2	AAW45039
63	15	100.0	31	2	AAW45039
64	15	100.0	31	2	AAW45039
65	15	100.0	31	2	AAW45039
66	15	100.0	31	2	AAW45039
67	15	100.0	31	2	AAW45039
68	15	100.0	31	2	AAW45039
69	15	100.0	31	2	AAW45039
70	15	100.0	31	2	AAW45039
71	15	100.0	31	2	AAW45039
72	15	100.0	31	2	AAW45039
73	15	100.0	31	2	AAW45039
74	15	100.0	31	2	AAW45039
75	15	100.0	31	2	AAW45039
76	15	100.0	31	2	AAW45039
77	15	100.0	31	2	AAW45039
78	15	100.0	31	2	AAW45039
79	15	100.0	31	2	AAW45039
80	15	100.0	31	2	AAW45039
81	15	100.0	31	2	AAW45039
82	15	100.0	31	2	AAW45039
83	15	100.0	31	2	AAW45039
84	15	100.0	31	2	AAW45039
85	15	100.0	31	2	AAW45039
86	15	100.0	31	2	AAW45039
87	15	100.0	31	2	AAW45039
88	15	100.0	31	2	AAW45039
89	15	100.0	31	2	AAW45039
90	15	100.0	31	2	AAW45039
91	15	100.0	31	2	AAW45039
92	15	100.0	31	2	AAW45039
93	15	100.0	31	2	AAW45039
94	15	100.0	31	2	AAW45039
95	15	100.0	31	2	AAW45039
96	15	100.0	31	2	AAW45039
97	15	100.0	31	2	AAW45039

	98	15	100.0	68	6	Adm59020 Propionib
	99	15	100.0	69	4	Aau17820 Novel hum
	100	15	100.0	69	7	Adg41200 Human res

ALIGNMENTS

RESULT 1

ADZ71692 ADZ71692 standard; peptide; 5 AA.

AC ADZ71692;

DT 14-JUL-2005 (first entry)

DE p21-derived peptide #277.

KM CDK inhibitor; Cyclin-dependent kinase-2 inhibitor; p21; cancer;

KW neoplasm; cytostatic; pharmaceutical; drug screening.

OS Synthetic.

PN WO2005040802-A2.

PD 06-MAY-2005.

PF 20-OCT-2004; 2004WO-GB004431.

PR 20-OCT-2003; 2003GB-00024466.

PR 02-FEB-2004; 2004US-00771242.

PA (CYCL-) CYCLACEL LTD.

PI Zheleva DI, Fischer PM, McInnes C, Andrews MJ1, Chan WC;

XX Atkinson GE;

XX WPI; 2005-355897/36.

PS Claim 20; Page 101; 112pp; English.

The invention relates to a peptide or its variant comprising formula: A-(B) m-C-(D) n-E, where m or n are each independently 0 or 1; A is a natural or unnatural amino acid residue having a side chain comprising at least one H-bond acceptor moiety and at least one H-bond donor moiety; each of B or D is independently an amino acid residue selected from arginine, glycine, citrulline, glutamine, serine, lysine, asparagine, isoleucine or alanine; C is a natural or unnatural amino acid residue having a branched or unbranched C 1-C 6 alkylene side chain optionally containing a H-bond donor or a H-bond acceptor moiety; and E is a natural or unnatural amino acid residue having an aryl or heteroaryl side chain. Also described are: a pharmaceutical composition comprising the peptide admixed with a diluent, an excipient or a carrier; an assay for identifying candidate substances capable of binding to a cyclin associated with a G1 control CDK enzyme and/or inhibiting the enzyme; an assay for identifying compounds that interact a cyclin or a cyclin when complexed with the physiologically relevant CDK; and a method of using a cyclin in a drug screening assay. The assay for identifying candidate substances capable of binding to a cyclin associated with a G1 control CDK enzyme and/or inhibiting the enzyme comprises: bringing into contact a peptide as defined above, the cyclin, the CDK and the candidate substance, under conditions where, in the absence of the candidate substance, being an inhibitor of interaction of the cyclin/CDK interaction, the peptidomimetic would bind to the cyclin; and monitoring any change in the expected binding of the peptide and the cyclin. The assay for identifying compounds that interact a cyclin or a cyclin when complexed with the physiologically relevant CDK comprises: incubating a candidate compound and the peptide and a cyclin or cyclin/CDK complex; and detecting binding of either the candidate compound or the peptide

with the cyclin. The cyclin is cyclin A, cyclin E or cyclin D. The assay comprises use of a three-dimensional model of a cyclin and a candidate compound. At least one of the assay components is bound to a solid phase. The peptidomimetic is labeled such as to emit a signal when bound to the cyclin. The cyclin is labeled such as to emit a signal when bound to the peptide. One of the assay components is labeled with a fluorescence emitter and the signal is detected using fluorescence polarization techniques. Using a cyclin in a drug screening assay comprises: selecting a candidate compound by performing rational drug design with a three-dimensional model of the cyclin, where the selecting is performed in conjunction with computer modeling; contacting the candidate compound with the cyclin; and detecting the binding of the candidate compound for the cyclin groove. A potential drug is selected on the basis of its having a greater affinity for the cyclin groove than that of the peptide. The method of detection comprises monitoring G0 and/or G1/S cell cycle, cell cycle-related apoptosis, suppression of E2F transcription factor, hypophosphorylation of cellular pRb, or in vitro anti-proliferative effects. The peptide is useful in preparing a medicament for treating a proliferative disorder, e.g., cancer. The present sequence represents a p21-derived peptide of the invention.

Sequence 5 AA;

Query-Match

Best Local Similarity 100.0%; Score 15; DB 9; Length 5; Pred. No. 2e+06; Mismatches 0; Indels 0; Gaps 0;

Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 4 RFLXF 8

1 RSLAF 5

RESULT 2

ADZ72055 ADZ72055 standard; peptide; 5 AA.

AC ADZ72055;

DT 14-JUL-2005 (first entry)

DE p21-derived peptide #640.

KM CDK inhibitor; Cyclin-dependent kinase-2 inhibitor; p21; cancer;

KW neoplasm; cytostatic; pharmaceutical; drug screening.

OS Synthetic.

PN WO2005040802-A2.

PD 06-MAY-2005.

PF 20-OCT-2004; 2004WO-GB004431.

PR 20-OCT-2003; 2003GB-00024466.

PR 02-FEB-2004; 2004US-00771242.

PA (CYCL-) CYCLACEL LTD.

PI Zheleva DI, Fischer PM, McInnes C, Andrews MJ1, Chan WC;

XX Atkinson GE;

XX WPI; 2005-355897/36.

PS Example 27; Page 83; 112pp; English.

The invention relates to a peptide or its variant comprising formula: A-(B) m-C-(D) n-E, where m or n are each independently 0 or 1; A is a natural or unnatural amino acid residue having a side chain comprising at least one H-bond acceptor moiety and at least one H-bond donor moiety;

DT 03-JUN-2004 (first entry)
XX Bcl-2 mutant CHK2 substrate domain.
DE kinase activity modulator; signaling enzyme; drug design; anti-tumour;
XX anti-inflammation; anti-ischaemia; bcl-2; CHK2; mutant; mutein.
KM
XX Unidentified.
OS Synthetic.
XX US2003170737-A1.
PN
XX 11-SEP-2003.
PD
XX 09-MAR-2002; 2002US-00093945.
PF
XX 09-MAR-2002; 2002US-00093945.
PR
XX (CARD/) CARDONE M H.
PA (YAF/) YAF/ M.
XX
PI Cardone M, Yaffe M;
XX WPI; 2003-787546/74.
DR
XX
XX A method for identifying a molecule capable of modulating a kinase
PT activity in situ and is useful in drug design to screen for molecules
PT that are candidates for anti-tumor, anti-inflammation, and anti-ischemia
PT therapy.
XX
PS Disclosure; Page 8; 25pp; English.
XX
XX The invention describes a method of identifying a molecule capable of
CC modulating a kinase activity in situ, comprising exposing a candidate
CC molecule to a cell comprising a signaling enzyme altered so as to bind a
CC phosphorylation substrate, where the substrate is associated with a
CC detectable label and the binding is regulated by a kinase, and
CC determining whether the candidate causes a change in expression of the
CC label. Also described are: a molecule capable of modulating a kinase
CC activity in situ identified by the claimed method; a fusion protein
CC comprising a genetically altered signaling enzyme and a label, where the
CC alteration produces an adapter molecule in the signaling enzyme capable
CC of binding to a phosphorylation substrate that the enzyme does not bind
CC in its unaltered state, where the binding is regulated by a kinase; an
CC isolated genetic molecule encoding the above fusion protein; a vector
CC capable of expressing the above genetic molecule; and a cell transfected
CC with the above vector. The invention is useful in drug design to screen
CC for molecules that are candidates for anti-tumour, anti-inflammation, and
CC anti-ischaemia therapy. This is the amino acid sequence of a Bcl-2 ERK1
CC substrate mutant that has been altered to change the kinase consensus
CC sequence from ERK1 to CHK2.
XX
SQ
Sequence 10 AA:
Query Match 100.0%; Score 15; DB 7; Length 10;
Best Local Similarity 60.0%; Pred. No. 3.1e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RXLXF 8
DB 4 RALSF 8

RESULT 5
ADE50790
ID ADE50790 standard; peptide; 11 AA.
XX
XX ADE50790;
AC
XX 29-JAN-2004 (first entry)
DT
XX
DE Protein kinase substrate polypeptide-related peptide 12.
XX

KM protein kinase substrate polypeptide; protein kinase activity;
KM threonine protein kinase; serine protein kinase.
XX
XX Unidentified.
OS Synthetic.
XX WO2003087400-A1.
PN
XX 23-OCT-2003.
PD
XX 26-MAR-2003; 2003WO-GB001286.
PF
XX 09-APR-2002; 2002GB-00008104.
PR
XX (UYDU-) UNIV DUNDEE.
PA
XX Armstrong C, Cohen P;
PI
XX WPI; 2003-845342/78.
DR
XX
XX New kit for assaying protein kinase activity and in screening for
PT modulators of protein kinase activity, comprises protein kinase substrate
PT polypeptides comprising a specificity conferring portion and a
PT phosphorylatable portion.
XX
XX
PS Disclosure; Fig 2; 48pp; English.
XX
XX This invention relates to a novel kit of parts comprising two or more
CC protein kinase substrate polypeptides. The kit and methods are useful in
CC assaying protein kinase activity and in screening for modulators of
CC protein kinase activity, in particular serine/threonine protein kinases.
CC The polypeptide or phosphorylated polypeptide or the antibody is also
CC used in an assay of protein kinase activity. The present sequence is that
CC of a peptide which is related to the invention.
XX
SQ
Sequence 11 AA:
Query Match 100.0%; Score 15; DB 7; Length 11;
Best Local Similarity 60.0%; Pred. No. 3.4e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RXLXF 8
DB 3 RLSLF 7

RESULT 6
ADY82927
ID ADY82927 standard; peptide; 11 AA.
XX
XX ADY82927;
AC
XX 02-JUN-2005 (first entry)
DT
XX
XX Protein kinase C (PKC) substrate peptide - SEQ ID 200.
DE
XX enzyme engineering; phosphorylation; protein kinase C; PKC; substrate.
KM
XX Synthetic.
OS
XX US2005064507-A1.
PN
XX 24-MAR-2005.
PD
XX 11-SEP-2003; 2003US-00660370.
PF
XX 11-SEP-2003; 2003US-00660370.
PR
XX 11-SEP-2003; 2003US-00660370.
PA (SHAW/) SHAW J S.
XX
XX Shaw JS;
PI
XX WPI; 2005-252669/26.
DR

Test set useful for characterizing substrate specificities of kinases, comprises two peptide pools, in which every peptide in each of the peptide pools has phosphorylatable, query, anchor and degenerate amino acid positions.

Claim 31; SEQ ID NO 200; 201pp; English.

The invention comprises a test set for characterizing substrate specificities of kinases, consisting of two peptide pools, in which every peptide in each of the peptide pools contains one phosphorylatable amino acid position, one query amino acid position, at least one anchor amino acid position, and at least one degenerate amino acid position. The test set of the invention is useful to determine the spectrum of peptide sequences that are phosphorylated by a kinase. The present amino acid sequence represents a peptide that was used as a substrate for a protein kinase C (PKC).

Sequence 11 AA;

Query Match Best Local Similarity 100.0%; Score 15; DB 9; Length 11; Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

4 RLLXF 8
5 RALLSF 9

RESULT 7
ADZ80906
ID ADZ80906 standard; peptide; 11 AA.
AC
XX ADZ80906;
XX
DT 14-JUL-2005 (first entry)
DE
XX Amino acid sequence of OAS1 gene conserved region #4.
XX
KW oligoadenylate synthetase gene; OAS1; viral infection;
KW hepatitis C virus infection; antiinflammatory; hepatotropic;
KW gastrointestinal disease; infection; cytostatic; varicella; antidiabetic;
KW endocrine disease; gastrointestinal disease; metabolic disorder;
KW neuroleptic; psychiatric disorder; neoplasm; andrology;
KW genitourinary disease; gene therapy; diabetes mellitus; schizophrenia;
KW cancer; prostate tumor; interferon-based treatment.
XX
OS Homo sapiens.
XX
PN WO2005040428-A2.
XX
PD 06-MAY-2005.
XX
PF 22-OCT-2004; 2004MO-US035284.
XX
PR 23-OCT-2003; 2003US-0513888P.
PR 06-FEB-2004; 2004US-0542373P.
PR 19-MAR-2004; 2004US-0554758P.
PR 08-APR-2004; 2004US-0560524P.
PR 09-JUN-2004; 2004US-0578323P.
PR 28-JUN-2004; 2004US-0583503P.
PR 26-AUG-2004; 2004US-0605243P.
XX
PA (ILLU-) ILLUMIGEN BIOSCIENCES INC.
XX
PI Iadonato SP, Magness CL, Rosenberg G, Scherer CA;
XX
XX WPI, 2005-333517/34.
XX
XX Identifying an oligoadenylate synthetase 1 gene mutation, useful for
XX treating, e.g. viral infection, cancer or diabetes, comprises detecting
XX in a nucleic acid sample, the presence of an oligoadenylate synthetase 1
XX point mutation.

xx Claim 3; SEQ ID NO 78; 11bp; English.

xx The specification describes a method for identifying an oligodeonyleate
CC synthetase gene (OAS1) mutation. The method comprises detecting in a
CC nucleic acid sample the presence of an OAS1 point mutation, preferably
CC substitution of a non-reference nucleotide for a reference nucleotide at
CC nucleotide position 2135728, 2135749, 2135978, 2144072, 2144088, 2144116,
CC 2144321, 2131025, 2133961, 2139587, 2144294, 2144985, 2156523, and
CC 2156632 of reference sequence ADZ80847. The mutations correlate with
CC resistance of humans to viral infection, such as Hepatitis C. The method
CC of the invention is useful for identifying an OAS1 mutation, and for
CC identifying susceptibility to viral infection, hepatitis C infection,
CC predisposition to diabetes mellitus or schizophrenia in humans,
CC susceptibility to cancer such as prostate cancer, or for identifying a
CC patient responsiveness to therapeutic treatments for viral infection,
CC where the therapeutic treatment is interferon-based and the patient
CC response is measured by sustained viral clearance. Peptides ADZ80903-
CC ADZ80912 represent conserved regions of the OAS1 gene.

SO Sequence 11 AA;

Query Match 100.0%; Score 15; DB 9; Length 11;
Best Local Similarity 60.0%; Pred. No. 3.4e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 RXLXF 8
| | | |
Db 3 RALSF 7

RESULT 8
ADY82979 standard; peptide; 12 AA.
ID ADY82979 standard; peptide; 12 AA.
XX ADY82979;
AC ADY82979;
XX 02-JUN-2005 (first entry)
DT 02-JUN-2005 (first entry)
XX Protein kinase C (PKC) substrate peptide - SEQ ID 321.
DE Protein kinase C (PKC) substrate peptide - SEQ ID 321.
XX enzyme engineering; phosphorylation; protein kinase C; PKC; substrate.
KW Synthetic.
XX OS
XX Key Location/Qualifiers
FH Misc-difference 8 /note= "Phosphorylated Thr"
FT
XX
PN US2005064507-A1.
XX
PD 24-MAR-2005.
PF 11-SEP-2003; 2003US-00660370.
XX 11-SEP-2003; 2003US-00660370.
PR 11-SEP-2003; 2003US-00660370.
XX (SHAW/) SHAW J S.
PA Shaw JS;
XX Shaw JS;
PI WPI; 2005-252669/26.
DR
XX Test set useful for characterizing substrate specificities of kinases,
PT comprises two peptide pools, in which every peptide in each of the
PT peptide pools has phosphorylatable, query, anchor and degenerate amino
PT acid positions.
XX
PS Claim 41; SEQ ID NO 321; 201pp; English.

The invention comprises a test set for characterizing substrate
specificities of kinases, consisting of two peptide pools, in which every
peptide in each of the peptide pools contains one phosphorylatable amino

CC acid position, one query amino acid position, at least one anchor amino
CC acid position, and at least one degenerate amino acid position. The test
CC set of the invention is useful to determine the spectrum of peptide
CC sequences that are phosphorylated by a kinase. The present amino acid
CC sequence represents a peptide that was used as a substrate for a protein
CC kinase C (PKC).

SO Sequence 12 AA;

Query Match 100.0%; Score 15; DB 9; Length 12;
Best Local Similarity 60.0%; Pred. No. 3.7e+02;

Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 RXLXF 8
DB 1 RLSPF 5

RESULT 9
ADE50791
ID ADE50791 standard; peptide; 13 AA.

XX ADE50791;

DT 29-JAN-2004 (first entry)

XX Protein kinase substrate polypeptide-related peptide 13.

XX Protein kinase substrate polypeptide; protein kinase activity;

KW threonine protein kinase; serine protein kinase.

XX Synthetic.

OS Undetermined.

XX WO2003087400-A1.

XX 23-OCT-2003.

XX 26-MAR-2003; 2003WO-GB001266.

XX 09-APR-2002; 2002GB-00008104.

XX (UYDU-) UNITV DUNDEE.

XX Armstrong C, Cohen P;

XX WPI; 2003-845342/78.

PT New kit for assaying protein kinase activity and in screening for
PT modulators of protein kinase activity, comprises protein kinase substrate
PT polypeptides comprising a specificity conferring portion and a
PT phosphorylatable portion.

PS Disclosure; Fig 2; 48pp; English.

CC This invention relates to a novel kit of parts comprising two or more
CC protein kinase substrate polypeptides. The kit and methods are useful in
CC assaying protein kinase activity and in screening for modulators of
CC protein kinase activity, in particular serine/threonine protein kinases.
CC The polypeptide or phosphorylated polypeptide or the antibody is also
CC used in an assay of protein kinase activity. The present sequence is that
CC of a peptide which is related to the invention.

SO Sequence 13 AA;

Query Match 100.0%; Score 15; DB 7; Length 13;
Best Local Similarity 60.0%; Pred. No. 4e+02;

Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 RXLXF 8
DB 5 RLSPF 9

RESULT 10
AAR30969
ID AAR30969 standard; peptide; 15 AA.

XX AAR30969;

DT 25-MAR-2003 (revised)

DT 07-MAY-1993 (first entry)

XX Jacalin fragment which interacts with CD4 receptor.

XX Jackfruit; human immunodeficiency virus; HIV-1; gp120; mitogenic lectin.

XX Artocarpus heterophyllus.

OS Key Location/Qualifiers

FT Misc-difference 1 /note= "uncharged hydrophilic amino acid - opt. absent"

XX WO9222574-A1.

XX 23-DEC-1992.

XX 05-JUN-1992; 92WO-FR000510.

XX 10-JUN-1991; 91FR-00007041.

XX 31-JAN-1992; 92FR-00001127.

XX (INRM) INSERM INST NAT SANTE & RECH MED.

XX Favero J, Dornand J, Corbeau P, Devaux C, Nicolas M, Liautard J;

XX WPI; 1993-018076/02.

XX Jacalin and its new peptide fragments for treating HIV - interacting with

XX the CD4 receptor and specifically preventing infection of lymphocytes.

XX Claim 7; Page 14; 27pp; French.

CC This fragment of Jacalin interacts with the CD4 receptor and is
CC homologous with the sequence of the HIV protein gp120. The peptide and
CC other peptides with biological activity equivalent to that of Jacalin are
CC useful in treatment of diseases caused by HIV. They specifically inhibit
CC infection of lymphocytes by HIV, do not affect normal lymphocyte function
CC and (unlike Jacalin itself) do not agglutinate cells. (Updated on 25-MAR-
CC 2003 to correct PN field.)

SO Sequence 15 AA;

Query Match 100.0%; Score 15; DB 2; Length 15;
Best Local Similarity 60.0%; Pred. No. 4.5e+02;

Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 RXLXF 8
DB 5 RLSPF 9

RESULT 11
ABP83159
ID ABP83159 standard; peptide; 17 AA.

XX ABP83159;

DT 04-MAR-2003 (first entry)

XX G protein-coupled receptor (GPCR) antigenic peptide SEQ ID NO:1832.

KW G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;
KW G protein-coupled receptor modulator; antibody; immune-related disease;
KW growth-related disease; cell regeneration-related disease; AIDS; cancer;
KW immunological-related cell proliferative disease; autoimmune disease;

KW Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;
 KW osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;
 KW graft versus host disease; Parkinson's disease; multiple sclerosis; pain;
 KW psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;
 KW mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;
 KW hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;
 KW ulcer.
 KM
 KM Homo sapiens.
 OS
 XX WO200261087-A2.
 PN
 XX 08-AUG-2002.
 PD
 XX 19-DEC-2001; 2001WO-US050107.
 PF
 XX 19-DEC-2000; 2000US-0257144P.
 PR
 XX (LIFE-) LIFESPAN BIOSCIENCES INC.
 PA
 XX Burner GC, Roush CL, Brown JP;
 PI
 XX WPI; 2003-046718/04.
 DR
 XX
 XX New isolated antigenic peptides e.g., for G protein-coupled receptors
 PT (GPCR), useful for diagnosing and designing drugs for treating conditions
 PT in which GPCRs are involved, e.g. AIDS, Alzheimer's disease, cancer or
 PT autoimmune diseases.
 XX
 XX Claim 1; Fig 2; 523pp; English.
 PS
 XX The present invention describes antigenic peptides (I) comprising: (a)
 CC any one of 1601 sequences (see ABP82019 to ABP83619) of 12-24 amino
 CC acids. Also described: (1) an assay for the detection of a particular G
 CC protein-coupled receptor (GPCR) or a candidate polypeptide in a sample;
 CC and (2) an isolated antibody having high specificity and high affinity or
 CC avidity for a particular GPCR. (I) can be used as GPCR modulators and in
 CC gene therapy. The antigenic peptides for GPCRs are useful in detecting an
 CC antibody against a particular GPCR, and in the production of specific
 CC antibodies. The peptides and antibodies are also useful for detecting the
 CC presence or absence of corresponding GPCRs. The antigenic peptides for
 CC GPCRs and antibodies are useful for diagnosing and designing drugs for
 CC treating immune-related diseases, growth-related diseases, cell
 CC regeneration-related diseases, immunological-related cell proliferative
 CC diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease,
 CC atherosclerosis, bacterial, fungal, protozoan or viral infections,
 CC osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute
 CC inflammation, allergies, Crohn's disease, diabetes, graft versus host
 CC disease, Parkinson's disease, multiple sclerosis, pain, psoriasis,
 CC anxiety, depression, schizophrenia, dementia, mental retardation, memory
 CC loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension,
 CC hypotension, renal disorders, rheumatoid arthritis, trauma, ulcers, or
 CC any other disorder in which GPCRs are involved. The antibodies may be
 CC used in immunoassays and immunodiagnosis. ABZ42523 to ABZ42869 encode
 CC GPCR proteins given in ABP81675 to ABP82018, which are used in the
 CC exemplification of the present invention
 CC
 XX
 XX Sequence 17 AA;
 SQ
 Query Match 100.0%; Score 15; DB 6; Length 17;
 Best Local Similarity 60.0%; Pred. No. 5, 1e+02;
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

XX
 DT 04-MAR-2003 (first entry)
 XX
 DE G protein-coupled receptor (GPCR) antigenic peptide SEQ ID NO:1838.
 XX
 KM G protein-coupled receptor (GPCR) antigenic peptide; gene therapy;
 KM G protein-coupled receptor modulator; antibody; immune-related disease;
 KM growth-related disease; cell regeneration-related disease; AIDS; cancer;
 KM immunological-related cell proliferative disease; autoimmune disease;
 KM Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;
 KM osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;
 KM graft versus host disease; Parkinson's disease; multiple sclerosis; pain;
 KM psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;
 KM mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;
 KM hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;
 KM ulcer.
 KM
 KM Homo sapiens.
 OS
 XX WO200261087-A2.
 PN
 XX 08-AUG-2002.
 PD
 XX 19-DEC-2001; 2001WO-US050107.
 PF
 XX 19-DEC-2000; 2000US-0257144P.
 PR
 XX (LIFE-) LIFESPAN BIOSCIENCES INC.
 PA
 XX Burner GC, Roush CL, Brown JP;
 PI
 XX WPI; 2003-046718/04.
 DR
 XX
 XX New isolated antigenic peptides e.g., for G protein-coupled receptors
 PT (GPCR), useful for diagnosing and designing drugs for treating conditions
 PT in which GPCRs are involved, e.g. AIDS, Alzheimer's disease, cancer or
 PT autoimmune diseases.
 XX
 XX Claim 1; Fig 2; 523pp; English.
 PS
 XX The present invention describes antigenic peptides (I) comprising: (a)
 CC any one of 1601 sequences (see ABP82019 to ABP83619) of 12-24 amino
 CC acids. Also described: (1) an assay for the detection of a particular G
 CC protein-coupled receptor (GPCR) or a candidate polypeptide in a sample;
 CC and (2) an isolated antibody having high specificity and high affinity or
 CC avidity for a particular GPCR. (I) can be used as GPCR modulators and in
 CC gene therapy. The antigenic peptides for GPCRs are useful in detecting an
 CC antibody against a particular GPCR, and in the production of specific
 CC antibodies. The peptides and antibodies are also useful for detecting the
 CC presence or absence of corresponding GPCRs. The antigenic peptides for
 CC GPCRs and antibodies are useful for diagnosing and designing drugs for
 CC treating immune-related diseases, growth-related diseases, cell
 CC regeneration-related disease, immunological-related cell proliferative
 CC diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease,
 CC atherosclerosis, bacterial, fungal, protozoan or viral infections,
 CC osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute
 CC inflammation, allergies, Crohn's disease, diabetes, graft versus host
 CC disease, Parkinson's disease, multiple sclerosis, pain, psoriasis,
 CC anxiety, depression, schizophrenia, dementia, mental retardation, memory
 CC loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension,
 CC hypotension, renal disorders, rheumatoid arthritis, trauma, ulcers, or
 CC any other disorder in which GPCRs are involved. The antibodies may be
 CC used in immunoassays and immunodiagnosis. ABZ42523 to ABZ42869 encode
 CC GPCR proteins given in ABP81675 to ABP82018, which are used in the
 CC exemplification of the present invention
 CC
 XX
 XX Sequence 18 AA;
 SQ
 Query Match 100.0%; Score 15; DB 6; Length 18;
 Best Local Similarity 60.0%; Pred. No. 5, 4e+02;
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 1 RLSLF 5

RESULT 13

ADR31537
ID ADR31537 standard; peptide; 19 AA.

AC ADR31537;

DT 04-NOV-2004 (first entry)

DE Bcl-2 mutant substrate peptide #1.

XX Drug designing; clinical application; anti-inflammatory; anti-tumour;
KW immune response; anti-ischaemic; atherosclerosis; therapy; Bcl-2;
KW mutant; muteln.

XX Unidentified.

OS

FT Key Location/Qualifiers

FT Misc-difference 2 /note= "wild type Thr substituted with Leu"

FT Misc-difference 4 /note= "wild type Pro substituted with Arg"

FT Misc-difference 8 /note= "wild type Pro substituted with Phe"

FT Misc-difference 11 /note= "wild type Ala substituted with Leu"

FT Misc-difference 13 /note= "wild type Pro substituted with Arg"

FT Misc-difference 17 /note= "wild type Pro substituted with Phe"

FT US2004157272-A1.

PD 12-AUG-2004.

XX 03-FEB-2004; 2004US-00771035.

XX 09-MAR-2002; 2002US-00093945.

XX (MERR-) MERRIMACK PHARM INC.

XX Cardone MR, Yaffe M;

XX WPI; 2004-580267/56.

XX Identifying molecule capable of modulating kinase activity in situ, by
PT exposing candidate molecule to cell comprising signaling enzyme altered
PT to bind substrate having label and determining whether molecule changes
PT expression of label.

XX Disclosure; Page 8; 25pp; English.

XX The invention relates to a method for identifying molecule capable of
CC modulating kinase activity in situ. The method involves exposing
CC candidate molecule to cell having signalling enzyme altered to bind
CC phosphorylation substrate associated with detectable label where the
CC binding between altered signalling enzyme and substrate is regulated by a
CC kinase and determining if the candidate molecule cause a change in
CC expression of label. The method is useful for testing and designing drugs
CC with various clinical application e.g. anti-inflammatory candidate
CC molecules. The invention is useful in the field of anti-tumour
CC therapeutics and immune response regulating drugs. The method is also
CC useful for developing anti-ischaemic drugs that are useful for treating
CC atherosclerosis. The present sequence is Bcl-2 mutant substrate peptide.
CC This sequence is used to illustrate the method of the invention.

XX Sequence 13: AA;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 4 RXLXF 8
Db 13 RLSLF 17

RESULT 14

ABJ10879
ID ABJ10879 standard; peptide; 20 AA.

AC ABJ10879;

DT 02-DEC-2002 (first entry)

DE Murine rhophilin protein SEQ ID NO: 164.

XX Human; PAL-18; cancer; chromosome 1q41; prostate cancer; colon cancer;
KW breast cancer; cytostatic.

XX Mus musculus.

XX US2002106765-A1.

XX 08-AUG-2002.

XX 12-MAR-2001; 2001US-00804682.

XX 10-MAR-2000; 2000US-0188586P.

XX (KIND/) KINDERS R J.

XX (CORE/) COREY M J.

XX Kinders RJ, Corey MJ;

XX WPI; 2002-697869/75.

XX New isolated PAL-18 polypeptide, useful for diagnosing, characterizing,
PT and treating disease and in determining disease susceptibility.

XX Disclosure; Page 105; 150pp; English.

XX The present invention relates to human PAL-18 polypeptides and
CC polynucleotides. The PAL-18 gene is found on chromosome 1q41. The
CC sequences can be used to diagnose, monitor and treat cancers, the
CC particularly breast, colon and prostate cancers. The present sequence is
CC a protein shown in the invention

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Query Match 100.0%; Score 15; DB 5; Length 20;
Best Local Similarity 60.0%; Pred. No. 6e+02; 2; Indels 0; Gaps 0;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RXLXF 8
Db 15 RALAF 19

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

G protein-coupled receptor (GPCR) antigenic peptide SEQ ID NO:1837.

G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;

G protein-coupled receptor modulator; antibody; immune-related disease;
growth-related disease; cell regeneration-related disease; AIDS; cancer;
immunological-related cell proliferative disease; autoimmune disease;

KW Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;
 KW osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;
 KW graft versus host disease; Parkinson's disease; multiple sclerosis; pain;
 KW psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;
 KW mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;
 KW hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;
 KW ulcer.
 KW
 XX Homo sapiens.
 OS
 PN WO200261087-A2.
 PD
 XX 08-AUG-2002.
 PF
 XX 19-DEC-2001; 2001WO-US050107.
 PR
 XX 19-DEC-2000; 2000US-0257144P.
 PA (LIFE-) LIFESPAN BIOSCIENCES INC.
 XX
 XX Burner GC, Roush CL, Brown JP;
 PI
 XX WPI; 2003-046718/04.
 DR
 XX
 PT New isolated antigenic peptides e.g., for G protein-coupled receptors
 PT (GPCR), useful for diagnosing and designing drugs for treating conditions
 PT in which GPCRs are involved, e.g. AIDS, Alzheimer's disease, cancer or
 PT autoimmune diseases.
 XX
 PS Claim 1; Fig 2; 523pp; English.
 CC
 CC The present invention describes antigenic peptides (1) comprising: (a)
 CC any one of 1601 sequences (see ABP82019 to ABP83619) of 12-24 amino
 CC acids. Also described: (1) an assay for the detection of a particular G
 CC protein-coupled receptor (GPCR) or a candidate polypeptide in a sample;
 CC and (2) an isolated antibody having high specificity and high affinity or
 CC avidity for a particular GPCR. (1) can be used as GPCR modulators and in
 CC gene therapy. The antigenic peptides for GPCRs are useful in detecting an
 CC antibody against a particular GPCR, and in the production of specific
 CC antibodies. The peptides and antibodies are also useful for detecting the
 CC presence or absence of corresponding GPCRs. The antigenic peptides for
 CC GPCRs and antibodies are useful for diagnosing and designing drugs for
 CC treating immune-related diseases, growth-related diseases, cell
 CC regeneration-related diseases, immunological-related cell proliferative
 CC diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease,
 CC atherosclerosis, bacterial, fungal, protozoan or viral infections,
 CC osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute
 CC inflammation, allergies, Crohn's disease, diabetes, graft versus host
 CC disease, Parkinson's disease, multiple sclerosis, pain, psoriasis,
 CC anxiety, depression, schizophrenia, dementia, mental retardation, memory
 CC loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension,
 CC hypotension, renal disorders, rheumatoid arthritis, trauma, ulcers, or
 CC any other disorder in which GPCRs are involved. The antibodies may be
 CC used in immunoassays and immunodiagnosis. ABZ42523 to ABZ42669 encode
 CC GPCR protein-given-in ABP81675 to ABP82018, which are used in the
 CC exemplification of the present invention
 CC
 XX
 SQ Sequence 20 AA;
 Query Match 100.0%; Score 15; DB 6; Length 20;
 Best Local Similarity 60.0%; Pred. No. 6e+02;
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 4 RXLXF 8
 DB 14 RSLSF 18
 RESULT 16
 ID ADM52410 standard; peptide; 20 AA.
 AC
 XX ADM52410;

XX
 DT 07-APR-2005 (first entry)
 XX
 DE Human PL protein PDZ binding peptide #425.
 XX
 KW antidepressant; cyostatic; uropathic; nocitropic; vulnerary; vasotrophic;
 KW neuroprotective; antidote; antiparkinsonian; ophthalmological;
 KW alpha-adrenergic receptor modulator; depression; apoptosis; migraine;
 KW Alzheimer's disease; Parkinson's disease; glaucoma; drug dependence.
 KW
 XX Homo sapiens.
 OS
 PN US2005019841-A1.
 PD
 XX 27-JAN-2005.
 PF
 XX 14-OCT-2003; 2003US-00684796.
 PR
 XX 14-MAY-1999; 99US-0134114P.
 PR 12-MAY-2000; 2000US-0057011B.
 PR 28-NOV-2000; 2000US-00724553.
 PR 16-FEB-2001; 2001US-0269523P.
 PR 03-AUG-2001; 2001US-0309841P.
 PR 19-FEB-2002; 2002US-00080273.
 PR 25-FEB-2002; 2002US-0360061P.
 PR 02-AUG-2002; 2002WO-US024655.
 PR 11-OCT-2002; 2002US-0418042P.
 PR 14-NOV-2002; 2002US-0426212P.
 XX
 PA (ARBO-) ARBOR VITA CORP.
 XX
 XX Garman JD, Lu PS;
 PI
 XX WPI; 2005-131132/14.
 DR
 XX
 PT New peptide that modulates binding between specific PDZ polypeptide and
 PT alpha-adrenergic receptor PL polypeptide, useful for treating depression,
 PT hypertrophy, Alzheimer's and Parkinson's.
 PT
 PS Example 3; Page 96; 11pp; English.
 CC
 CC The invention relates to a modulator (1) of binding between a specific
 CC PDZ polypeptide and an alpha-adrenergic receptor PL polypeptide, which is
 CC a peptide (P1) comprising at least 3 residues of a C-terminal sequence
 CC demonstrated to bind the target PDZ polypeptide, or a peptide mimetic
 CC (P2) of P1, or a small molecule having similar functional activity as the
 CC peptide with respect to PDZ polypeptide and PL polypeptide binding pair.
 CC Modulator of PDZ-PL interactions (claimed). (1) and P2 are useful for
 CC treating a disorder such as depression, lower urinary tract storage,
 CC prostate apoptosis, hypertrophy, proliferation and migration of vascular
 CC smooth muscle after carotid injury, migraine, coronary flow reserve
 CC following stenting, Alzheimer's, Parkinson's, neuroprotection, glaucoma
 CC and opioid withdrawal. This sequence corresponds to peptide from a PL
 CC protein-which binds a PDZ containing protein used in the invention.
 CC
 XX
 SQ Sequence 20 AA;
 Query Match 100.0%; Score 15; DB 9; Length 20;
 Best Local Similarity 60.0%; Pred. No. 6e+02;
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 4 RXLXF 8
 DB 1 RSLSF 5
 RESULT 17
 ID AAR51378 standard; protein; 21 AA.
 AC AAR51378;
 XX
 DT 25-MAR-2003 (reviled)

DT 20-OCT-1994 (first entry)
XX J alpha sequence (Val12.1/JaA?? usage).
DE
XX Rheumatoid arthritis; antibody; TCR; T cell receptor; lymphocyte;
KW expansion; complementary determining region; CDR3; antigen; MHC.
XX
OS Homo sapiens.
XX
PN WO9406823-A1.
XX
PD 31-MAR-1994.
XX
PF 14-SEP-1993; 93WO-US008644.
XX
PR 14-SEP-1992; 92US-00943418.
XX
PA (BGMN) BRIGHAM & WOMENS HOSPITAL.
XX
PI Brenner MB, Dersimonian H;
XX WPI; 1994-118395/14.
DR
XX Treatment and prevention of rheumatoid arthritis - using peptides derived
PT from DQw2 or antibodies to DQw2 which block activation of T lymphocytes.
XX
PS Disclosure; Fig 3A; 57pp; English.
XX
XX To gain insight into the basis for the Valpha12.1+T cell expansion in
CC rheumatoid arthritis, Valpha12.1 transcripts from positively selected
CC CD8+T cells were cloned and sequenced. In each of the three patients
CC analysed, distinct, repeated Valpha12.1 cong. sequences corresp. to
CC functional TCR alpha-chain transcripts were identified. All of the
CC repeated Valpha12.1+T cell rearrangements in the 3 patients analysed use
CC either Jalpa1a1, Jalpa1a2 or Jalpa1a6, each of which encodes a unique
CC sequence at the 3' end of the alpha gene segment. This short stretch of
CC shared residues (pro-tyr) is predicted to contribute (or is immediately
CC adjacent) to the third complementary determining region (CDR3) and thus
CC may play a role in antigen or MHC recognition. (Updated on 25-MAR-2003 to
CC correct PN field.)
XX
SQ Sequence 21 AA;
Query Match 100.0%; Score 15; DB 2; Length 21;
Best Local Similarity 60.0%; Pred. No. 6.3e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 4 RXLXF 8
DB 6 RALTF 10
RESULT 18
ABO60589
ID ABO60589 standard; protein; 21 AA.
XX
XX ABO60589;
AC
XX 29-JUL-2004 (first entry)
DT
XX
XX Human genome derived single exon protein #6823.
DE
XX
XX Human; gene expression; single exon probe; microarray;
KW alternative splicing event; genomic alteration.
XX
OS Homo sapiens.
XX
PN US2003194704-A1.
XX
PD 16-OCT-2003.
XX
PF 03-APR-2002; 2002US-00029386.
XX

ER 03-APR-2002; 2002US-00029386.
XX
XX (PENN/) PENN S G.
PA (RANK/) RANK D R.
PA (HANZ/) HANZEL D K.
XX
PI Penn SG, Rank DR, Hanzel DK.
XX
XX WPI; 2004-119264/12.
DR
XX
XX New human genome-derived single exon nucleic acid probes useful for human
PT gene expression analysis, for identifying or characterizing alternative
PT splicing events, for assessing genomic alterations or as tools for
PT surveying tissues.
XX
XX Claim 45; SEQ ID NO 34223; 80pp; English.
XX
XX The invention relates to a nucleic acid probe for measuring human gene
CC expression, comprising any of the 27,400 fully defined nucleotide
CC sequences in the specification, or their complements or fragments, and
CC encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC fully defined in the specification. The probe is a single exon probe that
CC hybridises under high stringency conditions to a nucleic acid molecule
CC expressed in human cells or tissues. Also included are a spatially-
CC addressable set of single exon nucleic acid probes for measuring human
CC gene expression (comprising a plurality of single exon nucleic acid
CC probes cited above, where each of the plurality of probes is separately
CC and addressably isolatable or amplifiable from the plurality), a single
CC exon microarray for measuring human gene expression, a method of
CC measuring human gene expression, a vector comprising the single exon
CC probe cited above, an ORF-encoded peptide comprising at least 8
CC contiguous amino acids of any of the above-mentioned amino acid
CC sequences (optionally with conservative amino acid substitutions), an
CC isolated antibody that binds specifically to a peptide cited above,
CC a method of selling and/or licensing single exon probes or microarrays to
CC a customer desiring to measure gene expression, a method of providing
CC human gene expression data by subscription, and a computer-readable
CC storage medium which contains a database having a plurality of records
CC (each record including data on the expression of a single exon probe
CC cited above). The probe, methods and apparatus are useful in gene
CC expression analysis. The probes may be used as tools for surveying
CC tissues to detect the presence of expressed messages that contain their
CC specific exon, or in constructing genome-derived single exon microarrays.
CC In addition, the probes are used in identifying and characterising
CC alternative splicing events, in detecting and characterising gross
CC alterations in the genomic locus that includes their exon, in assessing
CC smaller genomic alterations, in priming the synthesis of nucleic acids,
CC or in expressing the ORF-encoded peptide. The present sequence is a human
CC single exon probe protein of the invention. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?docID=20030194704
XX
SQ Sequence 21 AA;
Query Match 100.0%; Score 15; DB 8; Length 21;
Best Local Similarity 60.0%; Pred. No. 6.3e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 4 RXLXF 8
DB 9 RSLSF 13
RESULT 19
AAV78753
ID AAV78753 standard; peptide; 27 AA.
XX
XX AAV78753;
AC
XX 08-MAY-2000 (first entry)
DT
XX
XX Hypervariable region 1 representative peptide sequence 9.
DE

XX Hepatitis C virus; envelope protein E2; hypervariable region 1; minitope;
KW peptide library; treatment; prevent infection; antibody production.
XX
XX Hepatitis C virus.
XX
XX WO9960132-A1.
XX
XX 25-NOV-1999.
XX
XX 14-MAY-1999; 99WO-EP003344.
XX
XX 19-MAY-1998; 98GB-00010756.
XX
XX (RICE-) 1ST RICE RCHB BIOL MOLECOLARE ANGELETTI.
XX
XX PI Nicotia A, Lahm A, Tramontano A, Cortese R;
XX
XX WPI, 2000-126382/11.
XX
XX A new peptide library from hepatitis C virus, useful for production of
XX treatment for hepatitis C.
XX
XX Example; Page 73; 126pp; English.
XX
XX This sequence represents a peptide from the library of the invention. The
XX invention relates to a library of peptides which have an immunologically
XX reactive epitope of the hypervariable region 1 (HVR1) of envelope protein
XX 2 (E2) of hepatitis C virus. The peptides contained in the library
XX correspond to formulae given in the specification (see AAY78596-Y78598).
XX This sequence is included in a selection of a representative set of
XX natural HVR1 sequences. The peptides can be used in a method to select
XX antibodies which react with the HVR1 of E2 of hepatitis C virus, through
XX the selection of those antibodies which bind to the peptides. The
XX peptides from hepatitis C virus hypervariable region 1 of the envelope
XX protein E2 are used to produce a medicament for raising or increasing
XX levels of antibodies able to bind HCV (hepatitis C virus) HVR1 epitopes
XX in a mammal. The medicament is used to treat or prevent an HCV infection
XX
XX Sequence 27 AA;
XX
XX Query Match 100.0%; Score 15; DB 3; Length 27;
XX Best Local Similarity 60.0%; Pred. No. 8.1e+02;
XX Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 4 RXLXF 8
XX | | | |
XX 1 RTLSF 5
XX
XX Db
XX
XX RESULT 20
XX AAE18094
XX ID AAE18094 standard; peptide; 27 AA.
XX
XX AC AAE18094;
XX
XX DT 07-MAY-2002 (first entry)
XX
XX DE Biotinylated hepatitis C virus region derived peptide (HVR), 272.
XX
XX KW Hepatitis C virus; HCV conjugate; immune response; therapeutic; virucide;
XX hepatotropic; antiinflammatory; HCV region derived peptide; HVR.
XX
XX OS Hepatitis C virus.
XX
XX PN WO200193804-A2.
XX
XX PD 13-DEC-2001.
XX
XX PE 29-MAY-2001; 2001WO-US017302.
XX
XX PR 02-JUN-2000; 2000US-0209089P.
XX

PA (MERI) MERCK & CO INC.
XX
XX Conley AJ, McKenna PM, Przysiecki CT, Keller PM;
XX
XX WPI, 2002-164292/21.
XX
XX DR
XX
XX Hepatitis C virus conjugate useful for inducing immune response in a
XX subject comprising a polypeptide or protein complex carrier and
XX immunogenic peptides covalently bonded to the carrier.
XX
XX Example 1; Page 22; 63pp; English.
XX
XX The patent discloses hepatitis C virus (HCV) conjugates able to induce an
XX immune response recognising different strains and variants of HCV. The
XX conjugates comprise a polypeptide or protein complex carrier and one or
XX more HCV mimotopes. Sequences of the invention are useful for inducing an
XX immune response in a subject e.g. human, chimpanzees, mice or horses.
XX They are also useful for the preparation of antisera, in
XX therapeutic/diagnostic applications to generate anti-HCV antibodies, for
XX detecting the presence of HCV in a subject and treating the subject
XX infected with HCV. The present sequence is biotinylated HCV region
XX derived peptide (HVR), 270
XX
XX Sequence 27 AA;
XX
XX Query Match 100.0%; Score 15; DB 5; Length 27;
XX Best Local Similarity 60.0%; Pred. No. 8.1e+02;
XX Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 4 RXLXF 8
XX | | | |
XX 1 RTLSF 5
XX
XX Db
XX
XX RESULT 21
XX AAB38998
XX ID AAB38998 standard; peptide; 28 AA.
XX
XX AC AAB38998;
XX
XX DT 02-FEB-2001 (first entry)
XX
XX DE Human secreted peptide #20.
XX
XX KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
XX antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
XX vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
XX cardiac; gene therapy; cancer; immune disorder; cardiovascular disorder;
XX neurological disease; infection; human; secreted protein.
XX
XX OS Homo sapiens.
XX
XX PN WO200056880-A1.
XX
XX PD 28-SEP-2000.
XX
XX PF 16-MAR-2000; 2000WO-US006781.
XX
XX PR 19-MAR-1999; 99US-0125363P.
XX
XX PR 08-DEC-1999; 99US-0169617P.
XX
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX PI Rosen CA, Ruben SM, Komatsoulis G;
XX
XX WPI, 2000-602220/57.
XX
XX DR N-PSDB; AAC59706.
XX
XX PT Nucleic acid molecules encoding human secreted proteins, used in
XX preventing, treating or ameliorating disorders such as Parkinson's and
XX Alzheimer's diseases, cancers and infections.
XX
XX Claim 11; Page 380; 422pp; English.
XX

XX Sequences AAB38971-B39020 represent the amino acid sequences of 50 human
 CC secreted proteins encoded by the genes AAC59679-C59728. The genes and
 CC proteins are useful for preventing, ameliorating or treating medical
 CC conditions, e.g. by protein or gene therapy. The genes are isolated from
 CC a range of human tissues disclosed in the specification. The nucleic
 CC acids, proteins, antibodies and (ant)agonists are useful in the
 CC diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
 CC ovarian cancer, and other cancers of the adrenal gland, bone, bone
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b)
 CC immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic
 CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
 CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
 CC cardiovascular disorders such as myocardial ischaemia; (d) wound healing
 CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
 CC infectious diseases such as viral, bacterial, fungal and parasitic
 CC infections.

XX Sequence 28 AA;

SO Query Match 100.0%; Score 15; DB 3; Length 28;

Best Local Similarity 60.0%; Pred. No. 8.4e+02;

Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 RXLXF 8

DB 9 RFLAF 13

RESULT 22

AA02875 AAY02875 standard; protein; 30 AA.

AC AAY02875;

DT 11-JUN-1999 (first entry)

DE Fragment of human secreted protein encoded by gene 85.

XX Human; secreted protein; fusion protein; gene therapy; protein therapy;
 XX diagnosis; disease; cancer; tumour; neurodegenerative disorder; leukaemia;
 XX developmental abnormality; foetal deficiency; blood; allergy; renal;
 XX immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
 XX inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
 XX cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
 XX osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
 XX endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.

OS Homo sapiens.

PN WO9902546-A1.

PD 21-JAN-1999.

PF 07-JUL-1998; 98WO-US013684.

XX 08-JUL-1997; 97US-0051916P.

XX 08-JUL-1997; 97US-0051918P.

XX 08-JUL-1997; 97US-0051919P.

XX 08-JUL-1997; 97US-0051920P.

XX 08-JUL-1997; 97US-0051925P.

XX 08-JUL-1997; 97US-0051926P.

XX 08-JUL-1997; 97US-0051928P.

XX 08-JUL-1997; 97US-0051930P.

XX 08-JUL-1997; 97US-0051931P.

XX 08-JUL-1997; 97US-0052732P.

XX 08-JUL-1997; 97US-0052733P.

XX 08-JUL-1997; 97US-0052795P.

XX 08-JUL-1997; 97US-0052803P.

XX 18-AUG-1997; 97US-0055684P.

PR 18-AUG-1997; 97US-0055722P.
 PR 18-AUG-1997; 97US-0055723P.
 PR 18-AUG-1997; 97US-0055947P.
 PR 18-AUG-1997; 97US-0055948P.
 PR 18-AUG-1997; 97US-0055949P.
 PR 18-AUG-1997; 97US-0055950P.
 PR 18-AUG-1997; 97US-0055953P.
 PR 18-AUG-1997; 97US-0055954P.
 PR 18-AUG-1997; 97US-0055964P.
 PR 18-AUG-1997; 97US-0055984P.
 PR 18-AUG-1997; 97US-0056360P.
 PR 18-AUG-1997; 97US-0056360P.
 PR 12-SEP-1997; 97US-0058660P.
 PR 12-SEP-1997; 97US-0058661P.
 PR 12-SEP-1997; 97US-0058664P.
 PR 12-SEP-1997; 97US-0058785P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Fischer CL, Rosen CA, Soppet DR, Ruben SM, Kyaw H, Li Y, Zeng Z,
 PI Lafleur DW, Moore PA, Shi Y, Olsen HS, Ebner R, Brewer LA;

XX WPI; 1999-120770/10.

PT New isolated human genes and the secreted polypeptides they encode -
 PT useful for diagnosis and treatment of e.g. cancers, neurological
 PT disorders, immune diseases, inflammation or blood disorders.

PS Disclosure; Page 101; 464pp; English.

XX This sequence represents a fragment of a secreted human protein encoded
 CC by the nucleic acid molecule detailed in the descriptor line. The gene
 CC can be used to generate fusion proteins by linking to the gene to a human
 CC immunoglobulin Fc portion (e.g. AAX27302) for increasing the stability of
 CC the fused protein as compared to the human protein only. The invention
 CC relates to 123 novel genes and their fragments (nucleic acid sequences:
 CC AAX27311-X27449; amino acid sequences AAY02650-Y02788) which are useful
 CC for preventing, treating or ameliorating medical conditions e.g. by
 CC protein or gene therapy. Also, pathological conditions can be diagnosed
 CC by determining the amount of the new polypeptides in a sample or by
 CC determining the presence of mutations in the new polynucleotides.
 CC Specific uses are described for each of the 123 polynucleotides, based on
 CC which tissues they are most highly expressed in (see AAX27311 for
 CC described uses).

XX Sequence 30 AA;

SO Query Match 100.0%; Score 15; DB 2; Length 30;

Best Local Similarity 60.0%; Pred. No. 8.9e+02;

Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 RXLXF 8

DB 12 RFLTF 16

RESULT 23

ADA07705 ADA07705 standard; peptide; 30 AA.

AC ADA07705;

DT 06-NOV-2003 (first entry)

DE Human secreted protein from gene 85 peptide #5.

XX Immunosuppressive; dermatological; antiinflammatory; antiallergic;
 XX antiarthritic; human; autoimmune disease; autoimmune disorder; lupus;
 XX transplant rejection; allergic reaction; arthritis;
 XX squamous cell E48 antigen.

XX Homo sapiens.

OS US2003064412-A1.

XX

XX

XX

XX

XX 03-APR-2003.
 PD 30-OCT-2001; 2001US-00964440.
 XX 08-JUL-1997; 97US-0051916P.
 XX 08-JUL-1997; 97US-0051918P.
 PR 08-JUL-1997; 97US-0051919P.
 PR 08-JUL-1997; 97US-0051920P.
 PR 08-JUL-1997; 97US-0051925P.
 PR 08-JUL-1997; 97US-0051926P.
 PR 08-JUL-1997; 97US-0051928P.
 PR 08-JUL-1997; 97US-0051929P.
 PR 08-JUL-1997; 97US-0051930P.
 PR 08-JUL-1997; 97US-0051931P.
 PR 08-JUL-1997; 97US-0051932P.
 PR 08-JUL-1997; 97US-0052733P.
 PR 08-JUL-1997; 97US-0052734P.
 PR 08-JUL-1997; 97US-0052793P.
 PR 08-JUL-1997; 97US-0052795P.
 PR 08-JUL-1997; 97US-0052803P.
 PR 18-AUG-1997; 97US-0052804P.
 PR 18-AUG-1997; 97US-0052792P.
 PR 18-AUG-1997; 97US-0052723P.
 PR 18-AUG-1997; 97US-0055947P.
 PR 18-AUG-1997; 97US-0055948P.
 PR 18-AUG-1997; 97US-0055949P.
 PR 18-AUG-1997; 97US-0055950P.
 PR 18-AUG-1997; 97US-0055951P.
 PR 18-AUG-1997; 97US-0055953P.
 PR 18-AUG-1997; 97US-0055954P.
 PR 18-AUG-1997; 97US-0055964P.
 PR 18-AUG-1997; 97US-0055984P.
 PR 12-SEP-1997; 97US-0056360P.
 PR 12-SEP-1997; 97US-0058660P.
 PR 12-SEP-1997; 97US-0058661P.
 PR 12-SEP-1997; 97US-0058664P.
 PR 12-SEP-1997; 97US-0058785P.
 PR 07-JUL-1998; 98MO-US013684.
 PR 08-JAN-1999; 99US-00227357.
 XX (FISCHER/) FISCHER C. L.
 PA (ROSE/) ROSEN C. A.
 PA (SOPP/) SOPPET D. R.
 PA (RUBEN/) RUBEN S. M.
 PA (KYAW/) KYAW H.
 PA (LIYI/) LI Y.
 PA (ZENG/) ZENG Z.
 PA (LAFL/) LAFLEUR D. W.
 PA (MOOR/) MOORE P. A.
 PA (SHIY/) SHI Y.
 PA (OLSEN/) OLSEN H. S.
 PA (EBNER/) EBNER R.
 PA (BREW/) BREWER L. A.
 XX Fischer CL, Rosen CA, Soppet DR, Ruben SM, Kyaw H, Li Y, Zeng Z,
 PI Lafleur DW, Moore PA, Shi Y, Olsen HS, Ebner R, Brewer LA,
 XX WPI, 2003-540785/51.
 DR Novel antibody which specifically binds to a secreted protein useful for
 PT diagnosing and treating lupus, arthritis, allergic reactions, arthritis.
 XX
 XX Disclosure; Page 57; 355pp; English.
 PS
 XX The invention relates to an isolated antibody or its portion that
 CC specifically binds to a protein that shares sequence homology with human
 CC squamous cell E48 antigen, and consists of amino acid residues 21-116 or
 CC 1-116 the protein appearing as ADO7417 (one of 123 disclosed novel human
 CC secreted proteins encoded by 123 novel genes), or a protein consisting of
 CC amino acid sequence of secreted or full-length polypeptide encoded by
 CC HMFPO3 CDNA contained in ATCC Deposit No. 209126. The antibody is
 CC produced by immunising an animal with amino acid residues 21-116 of
 CC ADO7417, or with a protein consisting of amino acid sequence of the

CC secreted polypeptide encoded by the HMFPO3 CDNA contained in ATCC
 CC Deposit No. 209126, respectively. Also included are an isolated cell that
 CC produces the antibody and a hybridoma that produces the antibody. The
 CC antibody is a monoclonal, polyclonal, chimeric, humanised or human
 CC antibody. Optionally, the antibody is a Fab fragment, and is labeled by
 CC a label chosen from enzyme label, a radioisotope, and a fluorescent
 CC label. The antibody is useful as a probe for differential identification
 CC of tissues or cell types in which ADO7417 is expressed. The antibody is
 CC also for diagnosis and treatment of autoimmune diseases and disorders,
 CC such as lupus, transplant rejection, allergic reactions, and arthritis.
 CC The present sequence is a peptide/protein derived from one of the 123
 CC novel secreted proteins.
 XX
 SO Sequence 30 AA.
 Query Match 100.0%; Score 15; DB 7; Length 30;
 Best Local Similarity 60.0%; Pred. No. 8.9e+02;
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 4 RXLXF 8
 Db 12 RTLTF 16
 RESULT 24
 ID ADO741296 standard; peptide; 30 AA.
 AC ADO741296;
 XX
 DT 17-JUN-2004 (first entry)
 XX
 DE Novel human secreted protein fragment seqid 418.
 XX
 KW immunomodulator; immunosuppressive; antiinflammatory; dermatological;
 KW antiarthritic; antirheumatic; neuroprotective; antinaemic; muscular;
 KW antiallergic; antisthmatic; gastroprotective; anticoagulant;
 KW thrombolytic; antiarteriosclerotic; cardiact; cyostatic; nephrotropic;
 KW cardiovascular; respiratory; gene therapy; secreted protein;
 KW chromosome identification; hybrid mapping; gene expression control;
 KW immune system disorder; immunodeficiency; Chediak-Higashi syndrome;
 KW autoimmune disease; systemic lupus erythematosus; rheumatoid arthritis;
 KW multiple sclerosis; hemolytic anaemia; myasthenia gravis;
 KW allergic reaction; asthma; inflammatory condition;
 KW inflammatory bowel disease; B cell stimulator; T cell activator;
 KW blood-related disorder; eosinophilia; thrombosis; thromboembolism;
 KW atherosclerosis; myocardial infarction; angina; anaemia;
 KW hyperproliferative disorder; cancer; renal disorder;
 KW chronic kidney failure; renal tubular acidosis; kidney stone;
 KW cardiovascular disorder; respiratory disorder; human.
 XX
 OS Homo sapiens.
 XX
 PN US2004044191-A1.
 PD 04-MAR-2004.
 XX
 PF 10-OCT-2001; 2001US-00973278.
 XX
 PR 08-JUL-1997; 97US-0051916P.
 PR 08-JUL-1997; 97US-0051918P.
 PR 08-JUL-1997; 97US-0051919P.
 PR 08-JUL-1997; 97US-0051920P.
 PR 08-JUL-1997; 97US-0051925P.
 PR 08-JUL-1997; 97US-0051926P.
 PR 08-JUL-1997; 97US-0051928P.
 PR 08-JUL-1997; 97US-0051929P.
 PR 08-JUL-1997; 97US-0051930P.
 PR 08-JUL-1997; 97US-0051931P.
 PR 08-JUL-1997; 97US-0051932P.
 PR 08-JUL-1997; 97US-0052733P.
 PR 08-JUL-1997; 97US-0052734P.
 PR 08-JUL-1997; 97US-0052793P.
 PR 08-JUL-1997; 97US-0052795P.
 PR 08-JUL-1997; 97US-0052803P.
 PR 18-AUG-1997; 97US-0052804P.
 PR 18-AUG-1997; 97US-0052792P.
 PR 18-AUG-1997; 97US-0052723P.
 PR 18-AUG-1997; 97US-0055947P.
 PR 18-AUG-1997; 97US-0055948P.
 PR 18-AUG-1997; 97US-0055949P.
 PR 18-AUG-1997; 97US-0055950P.
 PR 18-AUG-1997; 97US-0055951P.
 PR 18-AUG-1997; 97US-0055953P.
 PR 18-AUG-1997; 97US-0055954P.
 PR 18-AUG-1997; 97US-0055964P.
 PR 18-AUG-1997; 97US-0055984P.
 PR 12-SEP-1997; 97US-0056360P.
 PR 12-SEP-1997; 97US-0058660P.
 PR 12-SEP-1997; 97US-0058661P.
 PR 12-SEP-1997; 97US-0058664P.
 PR 12-SEP-1997; 97US-0058785P.
 PR 07-JUL-1998; 97US-0058785P.
 PR 08-JAN-1999; 99US-00227357.

PR 08-JUL-1997; 97US-0052795P.
 PR 08-JUL-1997; 97US-0052803P.
 PR 18-AUG-1997; 97US-0055684P.
 PR 18-AUG-1997; 97US-0055722P.
 PR 18-AUG-1997; 97US-0055723P.
 PR 18-AUG-1997; 97US-0055947P.
 PR 18-AUG-1997; 97US-0055948P.
 PR 18-AUG-1997; 97US-0055949P.
 PR 18-AUG-1997; 97US-0055950P.
 PR 18-AUG-1997; 97US-0055953P.
 PR 18-AUG-1997; 97US-0055954P.
 PR 18-AUG-1997; 97US-0055964P.
 PR 18-AUG-1997; 97US-0055984P.
 PR 18-AUG-1997; 97US-0056360P.
 PR 12-SEP-1997; 97US-0058660P.
 PR 12-SEP-1997; 97US-0058661P.
 PR 12-SEP-1997; 97US-0058664P.
 PR 12-SEP-1997; 97US-0058785P.
 PR 07-JUL-1998; 98WO-US013684.
 PR 08-JAN-1999; 99US-00227357.
 PR 13-OCT-2000; 2000US-0239899P.
 XX
 PA (FISC/) FISCHER C. L.
 PA (ROSE/) ROSEN C. A.
 PA (SOBP/) SOPPET D. R.
 PA (RUBE/) RUBEN S. M.
 PA (KYAW/) KYAW H.
 PA (LIYY/) LI Y.
 PA (ZENG/) ZENG Z.
 PA (LAFU/) LAFLEUR D. W.
 PA (MOOR/) MOORE P. A.
 PA (SHIY/) SHI Y.
 PA (OLSE/) OLSEN H.
 PA (EBNE/) EBNER R.
 PA (BIRSE/) BIRSE C. E.
 XX
 PI Fischer CL, Rosen CA, Soppet DR, Ruben SM, Kyaw H, Li Y, Zeng Z,
 PI Lafleur DW, Moore PA, Shi Y, Olsen H, Ebner R, Birse CE;
 XX
 DR WPI: 2004-225733/21.
 XX
 PT New isolated nucleic acid encoding human proteins, useful for treating,
 PT preventing or diagnosing e.g. rheumatoid arthritis, multiple sclerosis,
 PT anemia, inflammatory bowel disease, atherosclerosis, cancers, chronic
 PT kidney failure.
 XX
 PS Disclosure; SEQ ID NO 418; 372pp; English.
 XX
 CC The invention describes novel human secreted proteins and the nucleotides
 CC encoding them. The polynucleotides are useful in chromosome
 CC identification, for radiation hybrid mapping, in controlling gene
 CC expression, in gene therapy or as molecular weight markers. The
 CC polynucleotides and polypeptides are useful for diagnosing, treating or
 CC preventing diseases of the immune system, immunodeficiencies, e.g.
 CC Chediak-Higashi syndrome, autoimmune diseases, e.g. systemic lupus
 CC erythematosus, rheumatoid arthritis, multiple sclerosis, haemolytic
 CC anaemia or myasthenia gravis, allergic reactions, e.g. asthma,
 CC inflammatory conditions, e.g. inflammatory bowel disease. They can also
 CC be used as a stimulator of B cell responsiveness to pathogens or as an
 CC activator of T cells. The polynucleotides and polypeptides are also
 CC useful for treating or preventing blood-related disorders, e.g.
 CC eosinophilia, thrombosis, thromboembolism, atherosclerosis, myocardial
 CC infarction, unstable angina or anaemia. They can also be used for
 CC treating, preventing or diagnosing hyperproliferative disorders
 CC (cancers), renal disorders (chronic kidney failure, renal tubular
 CC acidosis or kidney stones), cardiovascular disorders or respiratory
 CC disorders. This is the amino acid sequence of a novel human secreted
 CC protein fragment. Note: This sequence is available in electronic format
 CC from the US patent office at
 CC ftp.segdata.uspto.gov/sequence.html?DocID=20040044191.
 XX
 SQ Sequence 30 AA;

Query Match 100.0%; Score 15; DB 8; Length 30;
 Best Local Similarity 60.0%; Pred. No. 8.9e+02;
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Oy 4 RXLXF 8
 Db 12 RLTLF 16
 RESULT 25
 ID AAW45039 standard; peptide; 31 AA.
 AC AAW45039;
 DT 27-APR-1998 (first entry)
 DB Immunomodulatory peptide D22184AA.
 XX
 KW Immunomodulator; immunosuppressant; immunostimulator; treatment;
 KW transplant rejection; autoimmune disease; cancer; infection.
 XX
 OS Synthetic.
 PN WO9739023-A1.
 PD 23-OCT-1997.
 XX
 PE 04-APR-1997; 97WO-SE000574.
 XX
 PR 12-APR-1996; 96SE-00001422.
 PR 23-SEP-1996; 96SE-00003469.
 XX
 PA (ASTR) ASTRA AB.
 XX
 PI Bergstrand H, Eriksson T, Lindvall M, Saernstrand B;
 DR WPI: 1997-526397/48.
 DR N-PSDB; AAV05459.
 XX
 PT Nucleic acids encoding cysteine- or methionine-containing peptide(s)
 PT which have immuno-stimulatory or immunosuppressive activity - can be used
 PT to treat, e.g. cancers, infection, autoimmune disease or transplant
 PT rejection.
 XX
 PS Claim 22; Page 162; 183pp; English.
 XX
 CC The present peptide is an immunosuppressant or immunostimulator. An
 CC immunosuppressant can be used to treat transplant rejection or autoimmune
 CC disease, e.g. rheumatoid arthritis, systemic lupus erythematosus,
 CC Sjogren's syndrome, scleroderma, mixed connective tissue disease,
 CC dermatomyositis, polymyositis, Reiter's disease, Behcet's disease, type
 CC I diabetes, Hashimoto's thyroiditis, Graves' disease, multiple sclerosis,
 CC myasthenia gravis, encephalomyelitis, phemphigus vulgaris, vegetans or
 CC folliculitis, Senear-Usher syndrome or Brazilian phemphigus. An
 CC immunostimulator can be used to treat conditions such as cancer or
 CC infection
 XX
 SQ Sequence 31 AA;
 Query Match 100.0%; Score 15; DB 2; Length 31;
 Best Local Similarity 60.0%; Pred. No. 9.2e+02;
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Oy 4 RXLXF 8
 Db 23 RALAF 27
 RESULT 26
 ID AAV09486 standard; peptide; 31 AA.
 XX

AC AAY09486;
 XX 14-JUL-1999 (first entry)
 DT XX
 DE Immunoactive peptide containing a mammalian signal peptide #1.
 DE Immunoactive; immunomodulation; immunosuppression; immunostimulation;
 KM immune response; immunoreactive; autoimmune disease.
 XX
 OS Synthetic.
 XX WO9919347-A1.
 PN 22-APR-1999.
 XX
 PF 06-OCT-1998; 98WO-SE001801.
 XX
 PR 10-OCT-1997; 97US-00949024.
 XX
 PA (ASTR) ASTRA AB.
 PI Bergstrand H, Eriksson T, Lindvall M, Saernstrand B;
 XX WPI; 1999-287953/24.
 DR
 XX
 PT Synthetic genes encoding immunoreactive peptides containing cysteine or
 PT methionine.
 XX
 PS Disclosure; Page 12; 104pp; English.
 XX
 CC The present invention describes nucleic acid molecules comprising a
 CC coding sequence encoding an immunoreactive peptide and further encoding a
 CC protein targeting sequence. The nucleic acid is administered to a patient
 CC so that its expression product, an immunoreactive peptide, modulates an
 CC immune response in a patient. The nucleic acid can also be used to treat
 CC cancer, either after surgery to remove a portion of the cancer or after
 CC ionizing radiation. A cytokine is also administered in conjunction with
 CC the nucleic acid. Cells containing the nucleic acid molecule can also be
 CC used for treatment. The immunoreactive peptide is immunosuppressive and can
 CC be used in patients with autoimmune disease. The present sequence
 CC represents an immunoreactive peptide from the present invention
 XX
 SQ Sequence 31 AA;
 XX
 Quality Match 100.0%; Score 15; DB 2; Length 31;
 Best Local Similarity 60.0%; Pred. No. 9.2e+02;
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 4 RXLXF 8
 Db 23 RALAF 27
 XX
 RESULT 27
 ADV99762
 ID ADV99762 standard; peptide; 32 AA.
 XX
 AC ADV99762;
 XX
 DT 07-APR-2005 (first entry)
 XX
 DE Glucanase polypeptide signal sequence #32.
 XX
 KM Glucanase; hydrolysis; enzyme engineering; insect resistance;
 KM antibacterial; enzyme.
 XX
 OS Unidentified.
 XX
 PN WO2005003319-A2.
 XX
 PD 13-JAN-2005.
 XX
 PF 02-JUL-2004; 2004WO-US021492.

XX
 PR 02-JUL-2003; 2003US-0484725P.
 XX
 PA (DIVE-) DIVERSA CORP.
 XX
 PI Steer B, Callen W, Healey S, Pulliam D;
 XX WPI; 2005-112400/12.
 DR
 XX
 PT New isolated recombinant glucanase polypeptide, useful in e.g., pulp
 PT treatment, food processing, animal feeds, preparing dough, preparing fuel
 PT products, and in brewing.
 XX
 PS Disclosure; Page 171; 289pp; English.
 XX
 CC The invention relates to a glucanase polypeptide. The invention also
 CC relates to a nucleic acid encoding the polypeptide. The polypeptide is
 CC useful for hydrolyzing a cellulose (derivative) or a hemicellulose, e.g.,
 CC in a wood, paper pulp or a paper product, and for catalyzing hydrolysis
 CC of glucan in a feed, food product or a beverage. The feed, food product
 CC or beverage comprises a cereal-based animal feed, dough, wort or beer, or
 CC fruit or vegetable. The polypeptide is also useful for catalyzing
 CC hydrolysis of xylians in a microbe, plant cell, insect, yeast or mammalian
 CC cell, for catalyzing hydrolysis of a sugar to make a fuel product, for
 CC identifying its binding agents (e.g. modulators), and as a nutritional
 CC supplement in an animal diet, which involves preparing a nutritional
 CC supplement containing a glucanase enzyme comprising at least thirty
 CC contiguous amino acids of the glucanase polypeptide, and administering
 CC the nutritional supplement to an animal to increase utilization of xylan
 CC contained in feed or food ingested by the animal. Where the animal is a
 CC human, ruminant or a monogastric animal. The glucanase polypeptide is
 CC further useful for reducing lignin in a paper, wood or wood product, for
 CC eliminating or protecting animals from glucan containing microorganisms
 CC (especially Salmoneila), in oral care products, for producing or
 CC improving the flavor or texture of dairy products, and for producing
 CC small molecules. This sequence represents a glucanase polypeptide signal
 XX sequence-used in the scope of the invention.
 SQ Sequence 32 AA;
 XX
 Quality Match 100.0%; Score 15; DB 9; Length 32;
 Best Local Similarity 60.0%; Pred. No. 9.5e+02;
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 4 RXLXF 8
 Db 8 RLISF 12
 XX
 RESULT 28
 AAB44770
 ID AAB44770 standard; protein; 33 AA.
 XX
 AC AAB44770;
 XX
 DT 12-FEB-2001 (first entry)
 XX
 DE Human secreted protein sequence encoded by gene 9 SEQ ID NO:69.
 XX
 KM Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;
 KM antineumatic; antiproliferative; cytostatic; cardiant; vasotrophic;
 KM cerebroprotective; nootropic; neuroprotective; antibacterial; vinicude;
 KM fungicide; ophthalmological; gene therapy; autoimmune disease; infection;
 KM hyperproliferative disorder; cardiovascular disorder; angiogenesis;
 KM cerebrovascular disorder; nervous system disorder; ocular disorder;
 KM wound healing; skin aging; food additive; preservative.
 XX
 OS Homo sapiens.
 XX
 PN WO200058336-A1.
 XX
 PD 05-OCT-2000.
 XX

PF 23-MAR-2000; 2000WO-US007726.
 XX
 PR 26-MAR-1999; 99US-0126597P.
 PR 07-JAN-2000; 2000US-0174877P.
 XX
 BA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM, Komatsoulis G;
 XX
 DR WPI; 2000-602355/57.
 DR N-PSDB; AAC79807.
 XX
 PT Nucleic acid encoding human secreted proteins, used to treat, prevent,
 PT ameliorate or diagnose medical conditions such as cancer, and autoimmune
 PT diseases.
 XX
 PS Claim 11; Page 359; 391pp; English.
 XX
 CC The polynucleotide sequences given in AAC79799 to AAC79848 encode the
 CC human secreted proteins given in AAB44762 to AAB44811. AAB44812 to
 CC AAB44829 represent human secreted polypeptide sequences and proteins
 CC homologous to them, which are used in the exemplification of the present
 CC invention. Human secreted proteins have activities based on the tissues
 CC and cells the genes are expressed in. Examples of activities are:
 CC immunosuppressive; antiarthritic; antirheumatic; antiproliferative;
 CC cytostatic; cardiac; vasotropic; cerebroprotective; neurotropic;
 CC neuroprotective; antibacterial; virucide; fungicide; and
 CC ophthalmological. The polynucleotides and polypeptides can be used to
 CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
 CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
 CC in diagnosing a pathological condition or susceptibility to a
 CC pathological condition. Disorders which are diagnosed or treated include
 CC autoimmune diseases, hyperproliferative disorders, cardiovascular
 CC disorders, cerebrovascular disorders, angioneuromas, nervous system
 CC disorders, infections caused by bacteria, viruses and fungi and ocular
 CC disorders. The polypeptides can also be used to aid wound healing and
 CC epithelial cell proliferation, to prevent skin aging due to sunburn, to
 CC maintain organs before transplantation, for supporting cell culture of
 CC primary tissues, to regenerate tissues and in chemotaxis. The
 CC polypeptides can also be used as a food additive or preservative to
 CC increase or decrease storage capabilities. AAC79799 to AAC79798 and
 CC AAB44761 represent sequences used in the exemplification of the present
 CC invention.
 CC
 XX
 XX
 SQ Sequence 33 AA:
 Query Match 100.0%; Score 15; DB 3; Length 33;
 Best Local Similarity 60.0%; Pred. No. 9.8e+02;
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 4 RXLXF 8
 DB 6 RSLSF 10
 RESULT 29
 AAM32187
 ID AAM32187 standard; protein; 33 AA.
 XX
 AC AAM32187;
 XX
 DT 17-OCT-2001 (first entry)
 DE Peptide #6224 encoded by probe for measuring placental gene expression.
 XX
 DE Peptide #6224 encoded by probe for measuring placental gene expression.
 XX
 KW Probe; microarray; human; placenta; antenatal diagnosis;
 KW genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200157272-A2.
 XX
 PD 09-AUG-2001.

XX
 PF 30-JAN-2001; 2001WO-US000663.
 XX
 PR 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-488897/53.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for analyzing
 PT gene expression in human placenta.
 XX
 PS Claim 27; SEQ ID NO 32456; 654pp; English.
 XX
 CC The present invention relates to single exon nucleic acid probes (SENPs:
 CC see AAI31315-AA157546). The present sequence is a peptide encoded by one
 CC such probe. The probes are useful for producing a microarray for
 CC predicting, measuring and displaying gene expression in samples derived
 CC from human placenta. The probes are useful for antenatal diagnosis of
 CC human genetic disorders
 CC
 XX
 SQ Sequence 33 AA:
 Query Match 100.0%; Score 15; DB 4; Length 33;
 Best Local Similarity 60.0%; Pred. No. 9.8e+02;
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 4 RXLXF 8
 DB 14 RLSLF 18
 RESULT 30
 AAM60971
 ID AAM60971 standard; protein; 33 AA.
 XX
 AC AAM60971;
 XX
 DT 05-NOV-2001 (first entry)
 DE Human brain expressed single exon probe encoded protein SEQ ID NO: 33076.
 XX
 DE Human brain expressed single exon probe encoded protein SEQ ID NO: 33076.
 XX
 KW Human; brain expressed exon; gene expression analysis; probe; microarray;
 KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.
 XX
 OS Homo sapiens.
 XX
 PN WO200157275-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US000667.
 XX
 PR 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-483446/52.
 XX Single exon nucleic acid probes for analyzing gene expression in human
 PT brains.
 XX Example 4; SEQ ID NO 33076; 650bp + Sequence Listing; English.
 XX
 CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC brain. They can be used to measure gene expression in brain cell samples,
 CC which may enable the diagnosis and improved treatment of nervous system
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
 CC epilepsy and cancers. The present sequence is a protein encoded by one of
 CC the probes-of-the-invention
 XX
 SQ Sequence 33 AA;
 Query=Match 100.0%; Score 15; DB 4; Length 33;
 Best Local Similarity 60.0%; Pred. No. 9.8e+02;
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Oy 4 RXLXF 8
 Db 14 RTLSF 18
 RESULT 31
 ABG55401
 ID ABG55401 standard; peptide; 33 AA.
 XX
 AC ABG55401;
 XX
 DT 25-FEB-2003 (first entry)
 XX
 DE Human liver peptide, SEQ ID No 34049.
 XX
 KW Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
 KW hypercholesterolaemia; coronary heart disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200157273-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US000664.
 XX
 PR 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PT Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-488898/53.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for analyzing
 XX gene expression in human adult liver.
 XX
 PS Claim 27; SEQ ID NO 34049; 658bp; English.
 XX
 CC The invention relates to a single exon nucleic acid probe (SENP) (I) for
 CC measuring human gene expression in a sample derived from human adult
 CC liver comprising one of 13109 defined nucleotide sequences given in the
 CC specification (or complements/ fragments). The probe hybridises at high
 CC stringency to a nucleic acid molecule expressed in the human adult liver.
 CC (I) may be used for predicting, measuring and displaying gene expression
 CC in samples derived from human adult liver. The genes identified may be

CC involved in genetic liver diseases such as cirrhosis.
 CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
 CC associated with coronary heart disease. ABG47348-ABG59930 represent human
 CC liver single exon encoded peptides of the invention. Note: The sequence
 CC information for this patent does not appear in the printed specification
 CC but was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 33 AA;
 Query=Match 100.0%; Score 15; DB 4; Length 33;
 Best Local Similarity 60.0%; Pred. No. 9.8e+02;
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Oy 4 RXLXF 8
 Db 14 RTLSF 18
 RESULT 32
 ABG53589
 ID ABG53589 standard; peptide; 33 AA.
 XX
 AC ABG53589;
 XX
 DT 25-FEB-2003 (first entry)
 XX
 DE Human liver peptide, SEQ ID No 32237.
 XX
 KW Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
 KW hypercholesterolaemia; coronary heart disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200157273-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US000664.
 XX
 PR 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PT Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-488898/53.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for analyzing
 XX gene expression in human adult liver.
 XX
 PS Claim 27; SEQ ID NO 32237; 658bp; English.
 XX
 CC The invention relates to a single exon nucleic acid probe (SENP) (I) for
 CC measuring human gene expression in a sample derived from human adult
 CC liver comprising one of 13109 defined nucleotide sequences given in the
 CC specification (or complements/ fragments). The probe hybridises at high
 CC stringency to a nucleic acid molecule expressed in the human adult liver.
 CC (I) may be used for predicting, measuring and displaying gene expression
 CC in samples derived from human adult liver. The genes identified may be
 CC involved in genetic liver diseases such as cirrhosis.
 CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
 CC associated with coronary heart disease. ABG47348-ABG59930 represent human
 CC liver single exon encoded peptides of the invention. Note: The sequence
 CC information for this patent does not appear in the printed specification
 CC but was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX SO Sequence 33 AA; 100.0%; Score 15; DB 4; Length 33;
 Query Match: 100.0%; Score 15; DB 4; Length 33;
 Best Local Similarity: 60.0%; Pred. No. 9.8e+02;
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 RXLXF 8
 DB 14 RTLSF 18

RESULT 33
 ABG43538 standard; peptide; 33 AA.

XX ID ABG43538 standard; peptide; 33 AA.
 XX AC ABG43538;
 XX DT 19-AUG-2002 (first entry)
 XX DE Human peptide encoded by genome-derived single exon probe SEQ ID 33203.
 XX KM Human; single exon probe; asthma; lung cancer; COPD; ILD;
 XX KM chronic obstructive pulmonary disease; interstitial lung disease;
 XX KM familial idiopathic pulmonary fibrosis; neurofibromatosis;
 XX KM tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
 XX KM Heremansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
 XX KM pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
 XX KM pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 XX KM primary ciliary dyskinesia; pulmonary hypertension;
 XX KM hyaline membrane disease.

XX OS Homo sapiens.
 XX PN WO200186003-A2.
 XX PD 15-NOV-2001.
 XX PF 30-JAN-2001; 2001WO-US000665.
 XX PR 04-FEB-2000; 2000US-0180312P.
 XX PR 26-MAY-2000; 2000US-0207456P.
 XX PR 30-JUN-2000; 2000US-00608408.
 XX PR 03-AUG-2000; 2000US-00632366.
 XX PR 21-SEP-2000; 2000US-0234687P.
 XX PR 27-SEP-2000; 2000US-0236359P.
 XX PR 04-OCT-2000; 2000GB-00024263.
 XX PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX DR WPI; 2002-114183/15.
 XX PT Spatially-addressable set of single exon nucleic acid probes, used to
 XX measure gene expression in human lung samples.
 XX PS Claim 27; SEQ ID NO 33203; 634PP; English.
 XX CC The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of probes
 CC; the novel set of probes which hybridize at high stringency to a nucleic
 CC acid expressed in the human lung; measuring gene expression in a sample
 CC derived from human lung, comprising (a) contacting the array with a
 CC collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of the
 CC array; identifying exons in a eukaryotic genome, comprising (a)
 CC algorithmically predicting at least one exon from genomic sequences of
 CC the eukaryote; and (b) detecting specific hybridisation of detectably

CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene expression
 CC analysis, and for identifying exons in a gene, particularly using human
 CC lung derived mRNA and for the study of lung diseases such as asthma, lung
 CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
 CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
 CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Heremansky-
 CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
 CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
 CC Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
 CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The
 CC present sequence is a peptide/protein encoded by a single exon probe of
 CC the invention. Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic format
 CC directly from WFO at ftp.wipo.int/pub/published_pct_sequences

XX SO Sequence 33 AA; 100.0%; Score 15; DB 5; Length 33;
 Query Match: 100.0%; Score 15; DB 5; Length 33;
 Best Local Similarity: 60.0%; Pred. No. 9.8e+02;
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 RXLXF 8
 DB 14 RTLSF 18

RESULT 34
 ABG41720 standard; peptide; 33 AA.

XX ID ABG41720 standard; peptide; 33 AA.
 XX AC ABG41720;
 XX DT 19-AUG-2002 (first entry)
 XX DE Human peptide encoded by genome-derived single exon probe SEQ ID 31385.
 XX KM Human; single exon probe; asthma; lung cancer; COPD; ILD;
 XX KM chronic obstructive pulmonary disease; interstitial lung disease;
 XX KM familial idiopathic pulmonary fibrosis; neurofibromatosis;
 XX KM tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
 XX KM Heremansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
 XX KM pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
 XX KM pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 XX KM primary ciliary dyskinesia; pulmonary hypertension;
 XX KM hyaline membrane disease.

XX OS Homo sapiens.
 XX PN WO200186003-A2.
 XX PD 15-NOV-2001.
 XX PF 30-JAN-2001; 2001WO-US000665.
 XX PR 04-FEB-2000; 2000US-0180312P.
 XX PR 26-MAY-2000; 2000US-0207456P.
 XX PR 30-JUN-2000; 2000US-00608408.
 XX PR 03-AUG-2000; 2000US-00632366.
 XX PR 21-SEP-2000; 2000US-0234687P.
 XX PR 27-SEP-2000; 2000US-0236359P.
 XX PR 04-OCT-2000; 2000GB-00024263.
 XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX 03-APR-2002; 2002US-00029386.
 PF 03-APR-2002; 2002US-00029386.
 PR (PENN/) PENN S G.
 XX (RANK/) RANK D R.
 PA (HANZ/) HANZEL D K.
 XX Penn SG, Rank DR, Hanzel DK;
 PI WPI: 2004-119264/12.
 DR
 XX New human genome-derived single exon nucleic acid probes useful for human
 PT gene expression analysis, for identifying or characterizing alternative
 PT splicing events, for assessing genomic alterations or as tools for
 PT surveying tissues.
 PF
 XX Claim 45; SEQ ID NO 30151; 80bp; English.
 PS
 XX The invention relates to a nucleic acid probe for measuring human gene
 CC expression, comprising any of the 27,400 fully defined nucleotide
 CC sequences in the specification, or their complements or fragments, and
 CC encoding at least 8 amino acids or any of the 688 amino acid sequences
 CC fully defined in the specification. The probe is a single exon probe that
 CC hybridizes under high stringency conditions to a nucleic acid molecule
 CC expressed in human cells or tissues. Also included are a spatially-
 CC addressable set of single exon nucleic acid probes for measuring human
 CC gene expression (comprising a plurality of single exon nucleic acid
 CC probes cited above, where each of the plurality of probes is separately
 CC and addressably isolatable or amplifiable from the plurality), a single
 CC exon microarray for measuring human gene expression, a method of
 CC measuring human gene expression, a vector comprising the single exon
 CC probe cited above, an ORF-encoded peptide comprising at least 8
 CC contiguous amino acids of any of the above-mentioned amino acid
 CC sequences (optionally with conservative amino acid substitutions), an
 CC isolated antibody that binds specifically to a peptide cited above,
 CC a method of selling and/or licensing single exon probes or microarrays to
 CC a customer desiring to measure gene expression, a method of providing
 CC human gene expression data by subscription, and a computer-readable
 CC storage medium which contains a database having a plurality of records
 CC (each record including data on the expression of a single exon probe
 CC cited above. The probe, methods and apparatus are useful in gene
 CC expression analysis. The probes may be used as tools for surveying
 CC tissues to detect the presence of expressed messages that contain their
 CC specific exon, or in constructing genome-derived single exon microarrays.
 CC In addition, the probes are used in identifying and characterizing
 CC alternative splicing events, in detecting and characterizing gross
 CC alterations in the genomic locus that includes their exon, in assessing
 CC smaller genomic alterations, in priming the synthesis of nucleic acids,
 CC or in expressing the ORF-encoded peptide. The present sequence is a human
 CC single exon probe protein of the invention. Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from USPTO at
 CC seqdata.uspto.gov/sequence.html?DocID=20030194704
 XX
 XX Sequence 35 AA:
 QY
 DB 4 RLTXF 8
 18 RALAF 22
 Query Match 100.0%; Score 15; DB 8; Length 35;
 Best Local Similarity 60.0%; Pred. No. 1e+03;
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 RESULT 37
 AAB63066
 ID AAB63066 standard; protein; 36 AA.
 XX
 AC AAB63066;
 XX

DT 26-MAR-2001 (first entry)
 XX
 DE Human secreted protein sequence encoded by gene 18 SEQ ID NO:76.
 XX
 KW Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;
 KW antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic;
 KW cerebroprotective; neuroprotective; antibacterial; virucide;
 KW fungicide; ophthalmological; vulnery; gene therapy; neoplasm;
 KW autoimmune disease; rheumatoid arthritis; hyperproliferative disorder;
 KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
 KW cerebral ischaemia; angiogenesis; nervous system disorder; infection;
 KW Alzheimer's disease; ocular disorder; corneal infection; wound healing;
 KW skin aging; food additive; preservative.
 XX
 OS Homo sapiens.
 XX
 PN WO200061748-A1.
 XX
 PD 19-OCT-2000.
 XX
 PF 06-APR-2000; 2000WO-US008982.
 XX
 PR 09-APR-1999; 99US-0128696P.
 XX 14-JAN-2000; 2000US-0176069P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM, Komatsoulis G;
 XX WPI: 2000-638566/61.
 XX N-PSDB; AAF22333.
 DR
 XX New nucleic acid molecules encoding 48 human secreted proteins for
 PT diagnosing, preventing, treating or ameliorating medical conditions and
 PT used as food additives or preservatives.
 PF
 XX Claim 11; Page 436; 480bp; English.
 PS
 XX AAF22316 to AAF22363 encode the human secreted proteins given in AAB63049
 CC to AAB63096. AAB63097 to AAB63113 represent more human secreted proteins
 CC and polypeptides homologous to them. Human secreted proteins have
 CC activities based on the tissues and cells the genes are expressed in.
 CC Examples of activities include: immunosuppressive; antiarthritic;
 CC antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic;
 CC cerebroprotective; neuroprotective; antibacterial; virucide;
 CC fungicide; ophthalmological; and vulnery. The polynucleotides and
 CC proteins can be used to prevent, treat or ameliorate a medical condition
 CC in e.g. humans, mice, rabbits, goats, horses, cats, dogs, chickens or
 CC sheep. They are also used in diagnosing a pathological condition or
 CC susceptibility to a pathological condition. Disorders which are diagnosed
 CC or treated include autoimmune diseases e.g. rheumatoid arthritis,
 CC hyperproliferative disorders e.g. neoplasms of the breast or liver,
 CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders
 CC e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g.
 CC Alzheimer's disease, infections caused by bacteria, viruses and fungi and
 CC ocular disorders e.g. corneal infection. The polypeptides can also be
 CC used to aid wound healing and epithelial cell proliferation, to prevent
 CC skin aging due to sunburn, to maintain organs before transplantation, for
 CC supporting cell culture of primary tissues, to regenerate tissues and in
 CC chemotaxis. The polypeptides can also be used as a food additive or
 CC preservative to increase or decrease storage capabilities, fat content,
 CC lipid, protein, carbohydrate, vitamins, minerals, cofactors and other
 CC nutritional components. AAF22307 to AAF22315 and AAB63048 represent
 CC sequences used in the exemplification of the present invention
 XX
 XX Sequence 36 AA:
 QY
 DB 4 RLTXF 8
 Query Match 100.0%; Score 15; DB 3; Length 36;
 Best Local Similarity 60.0%; Pred. No. 1.1e+03;
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 RESULT 37
 AAB63066
 ID AAB63066 standard; protein; 36 AA.
 XX
 AC AAB63066;
 XX

DB 31 RSLAF 35

RESULT 38
ABG68830
ID ABG68830 standard; protein; 38 AA.
XX
AC ABG68830;
XX
DT 07-OCT-2002 (first entry)
XX
DE Cytochrome P450 3A7 (CYP3A7) variant C-terminal ORF #2.
XX
KW Cytochrome P450; CYP3A1; CYP3A2; CYP3A3; CYP3A4; CYP3A5; CYP3A7;
KM drug metabolism; drug design; drug screening.
XX
OS Synthetic.
XX
PN MO200244213-A1.
XX
PD 06-JUN-2002.
XX
PF 28-NOV-2001; 2001WO-SE002631.
XX
PR 28-NOV-2000; 2000SE-00004366.
XX
PR 11-JUN-2001; 2001SE-00002061.
XX
PA (ZAPH/) ZAPHIROPOULOS P G.
XX
PA (FINNT/) FINNTA C.
XX
PI Zaphiropoulos PG, Finnta C;
XX
DR WPI; 2002-557532/59.
XX
PT Novel cytochrome P450 protein in which CYP3A43 exon 1 is joined to sets
PT of CYP3A4 or CYP3A5 exons, useful as medicament, and in evaluating drug
PT metabolism, in drug design and drug screening.
XX
PS Disclosure; Fig 2; 131pp; English.
XX
CC The invention describes a cytochrome P450 protein (I) in which CYP3A43
CC exon 1 is joined to sets of CYP3A4 or CYP3A5 exons, as well as sub
CC fragments, variants and multiples of (I) having essentially the same
CC characteristics. (I) is useful as a medicament, and for evaluating drug
CC metabolism, in drug design, and drug screening, and in tests for
CC adjusting the dose of drugs. This is the amino acid sequence of a
CC cytochrome P450 3A7 (CYP3A7) variant C-terminal
XX
SQ Sequence 38 AA;
Query Match 100.0%; Score 15; DB 5; Length 38;
Best Local Similarity 60.0%; Pred. No. 1.1e+03;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 RXLXF 8
DB 15 RSLAF 19

RESULT 39
AAG15201
ID AAG15201 standard; protein; 40 AA.
XX
AC AAG15201;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 15360.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.

OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
PF 25-FEB-1999; 99US-0121825P.
XX
PR 05-MAR-1999; 99US-0123180P.
XX
PR 09-MAR-1999; 99US-0123548P.
XX
PR 23-MAR-1999; 99US-0125788P.
XX
PR 25-MAR-1999; 99US-0126254P.
XX
PR 29-MAR-1999; 99US-0126785P.
XX
PR 01-APR-1999; 99US-0127462P.
XX
PR 06-APR-1999; 99US-0128234P.
XX
PR 08-APR-1999; 99US-0128714P.
XX
PR 16-APR-1999; 99US-0129845P.
XX
PR 19-APR-1999; 99US-0130077P.
XX
PR 21-APR-1999; 99US-0130449P.
XX
PR 23-APR-1999; 99US-0130510P.
XX
PR 28-APR-1999; 99US-0130891P.
XX
PR 30-APR-1999; 99US-0132048P.
XX
PR 04-MAY-1999; 99US-0132407P.
XX
PR 05-MAY-1999; 99US-0132485P.
XX
PR 06-MAY-1999; 99US-0132486P.
XX
PR 06-MAY-1999; 99US-0132487P.
XX
PR 07-MAY-1999; 99US-0132863P.
XX
PR 11-MAY-1999; 99US-0134256P.
XX
PR 14-MAY-1999; 99US-0134218P.
XX
PR 14-MAY-1999; 99US-0134219P.
XX
PR 14-MAY-1999; 99US-0134221P.
XX
PR 14-MAY-1999; 99US-0134370P.
XX
PR 18-MAY-1999; 99US-0134768P.
XX
PR 19-MAY-1999; 99US-0134941P.
XX
PR 20-MAY-1999; 99US-0135124P.
XX
PR 21-MAY-1999; 99US-0135353P.
XX
PR 24-MAY-1999; 99US-0135629P.
XX
PR 25-MAY-1999; 99US-0136021P.
XX
PR 27-MAY-1999; 99US-0136392P.
XX
PR 28-MAY-1999; 99US-0136782P.
XX
PR 01-JUN-1999; 99US-0137222P.
XX
PR 03-JUN-1999; 99US-0137528P.
XX
PR 04-JUN-1999; 99US-0137502P.
XX
PR 07-JUN-1999; 99US-0137724P.
XX
PR 08-JUN-1999; 99US-0138094P.
XX
PR 10-JUN-1999; 99US-0138540P.
XX
PR 10-JUN-1999; 99US-0138847P.
XX
PR 14-JUN-1999; 99US-0139119P.
XX
PR 16-JUN-1999; 99US-0139452P.
XX
PR 16-JUN-1999; 99US-0139453P.
XX
PR 17-JUN-1999; 99US-0139492P.
XX
PR 18-JUN-1999; 99US-0139454P.
XX
PR 18-JUN-1999; 99US-0139455P.
XX
PR 18-JUN-1999; 99US-0139456P.
XX
PR 18-JUN-1999; 99US-0139457P.
XX
PR 18-JUN-1999; 99US-0139458P.
XX
PR 18-JUN-1999; 99US-0139459P.
XX
PR 18-JUN-1999; 99US-0139460P.
XX
PR 18-JUN-1999; 99US-0139461P.
XX
PR 18-JUN-1999; 99US-0139462P.
XX
PR 18-JUN-1999; 99US-0139463P.
XX
PR 18-JUN-1999; 99US-0139750P.
XX
PR 18-JUN-1999; 99US-0139763P.
XX
PR 21-JUN-1999; 99US-0139817P.
XX
PR 22-JUN-1999; 99US-0139889P.
XX
PR 23-JUN-1999; 99US-0140353P.
XX
PR 23-JUN-1999; 99US-0140354P.
XX
PR 24-JUN-1999; 99US-0140689P.
XX
PR 28-JUN-1999; 99US-0140823P.
XX
PR 29-JUN-1999; 99US-0140991P.

PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0144390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 19-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 20-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 21-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145195P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145244P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145313P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145819P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146388P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 06-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 09-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-015138P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.

PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155486P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-015753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 100.0%; Score 15; DB 3; Length 40;
Best Local Similarity 60.0%; Pred. No. 1.2e+03;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXP 8
Db 12 RSLSP 16

RESULT 40
ADD90505
ID ADD90505 standard; protein; 40 AA.
XX
AC ADD90505;
XX
DT 29-JAN-2004 (first entry)
XX
DE Novel human secreted protein seq id 68 protein feature seq id 320.
XX gene therapy; cytosstatic; cancer; human; secreted protein.
XX
OS Homo sapiens.
XX
PN US2003199683-A1.
XX
PD 23-OCT-2003.
XX
PF 30-MAR-2001; 2001US-00820649.
XX
PR 30-JUL-1997; 97US-0054209P.

```

PR      30-JUL-1997;    97US-0054211P.
PR      30-JUL-1997;    97US-0054212P.
PR      30-JUL-1997;    97US-0054213P.
PR      30-JUL-1997;    97US-0054214P.
PR      30-JUL-1997;    97US-0054215P.
PR      30-JUL-1997;    97US-0054216P.
PR      30-JUL-1997;    97US-0054217P.
PR      30-JUL-1997;    97US-0054218P.
PR      30-JUL-1997;    97US-0054234P.
PR      30-JUL-1997;    97US-0054236P.
PR      18-AUG-1997;    97US-0055968P.
PR      18-AUG-1997;    97US-0055969P.
PR      18-AUG-1997;    97US-0055972P.
PR      19-AUG-1997;    97US-0056534P.
PR      19-AUG-1997;    97US-0056543P.
PR      19-AUG-1997;    97US-0056544P.
PR      19-AUG-1997;    97US-0056561P.
PR      19-AUG-1997;    97US-0056727P.
PR      19-AUG-1997;    97US-0056729P.
PR      19-AUG-1997;    97US-0056730P.
PR      26-JUN-1998;    98WO-US015949.
PR      21-SEP-2000;    2000US-00666987.
XX
XX      (RUBE/) RUBEN S M.
PA      (FENG/) FENG P.
PA      (LAFI/) LAFLEUR D W.
PA      (MOOR/) MOORE P A.
PA      (SHIY) SHI Y.
PA      (KYAW) KYAW H.
PA      (LIY/) LI Y.
PA      (ZENG/) ZENG Z.
PA      (CART/) CARTER K C.
PA      (ENDR/) ENDRESS G A.
PA      (WEIY/) WEI Y.
PA      (PANP/) PAN P.
PA      (ROSE/) ROSEN C A.
XX
XX      Ruben SM, Feng P, Lafleur DW, Moore PA, Shi Y, Kyaw H, Li Y;
PI      Zeng Z, Carter KC, Endress GA, Wei Y, Fan P, Rosen CA;
DR      WPI; 2003-852813/79.
XX
XX      New nucleic acid molecule, useful for preparing a medicament for
PT      preventing, treating or ameliorating a medical condition e.g., cancer.
XX
XX      Disclosure: SEQ ID NO 320; 213pp; English.
PS
XX
XX      The invention describes novel isolated human nucleic acids. The nucleic
CC      acid is useful for preparing a medicament for preventing, treating or
CC      ameliorating a medical condition e.g., cancer, and in gene therapy. This
CC      is the amino acid sequence of polypeptide feature of a novel human
CC      secreted protein of the invention.
XX
XX      Sequence 40 AA:
SO
Oy      4 RXLXF 8
Db      2 RLSLF 6
Query Match          100.0%; Score 15; DB 7; Length 40;
Bee Local Similarity 60.0%; Pred.No.1.2e+03;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

XX	Secreted protein; gene therapy; neural disorder; immune system disorders;
KM	muscular disorder; reproductive disorder; gastrointestinal disorder;
KM	pulmonary disorder; cardiovascular disorder; renal disorder;
KM	proliferative disorder; cancer; systemic lupus erythematosus;
KM	rheumatoid arthritis; multiple sclerosis; thyroiditis; anaemia;
KM	Graves's disease; diabetes; hepatitis; asthma; allergy; nephritis;
KM	Parkinson's disease; Alzheimer's disease; atherosclerosis;
KM	myocardial infarction; AIDS; infection; human.
OS	
OS	Homo sapiens.
PM	US2003166541-A1.
PD	
XX	04-SEP-2003.
PF	
XX	04-JUN-2002; 2002US-00160162.
XX	
PR	30-JUL-1997; 97US-0054209P.
PR	30-JUL-1997; 97US-0054211P.
PR	30-JUL-1997; 97US-0054212P.
PR	30-JUL-1997; 97US-0054213P.
PR	30-JUL-1997; 97US-0054214P.
PR	30-JUL-1997; 97US-0054215P.
PR	30-JUL-1997; 97US-0054217P.
PR	30-JUL-1997; 97US-0054218P.
PR	30-JUL-1997; 97US-0054234P.
PR	30-JUL-1997; 97US-0054236P.
PR	18-AUG-1997; 97US-0055968P.
PR	18-AUG-1997; 97US-0055969P.
PR	18-AUG-1997; 97US-0055972P.
PR	19-AUG-1997; 97US-0056534P.
PR	19-AUG-1997; 97US-0056543P.
PR	19-AUG-1997; 97US-0056544P.
PR	19-AUG-1997; 97US-0056561P.
PR	19-AUG-1997; 97US-0056727P.
PR	19-AUG-1997; 97US-0056729P.
PR	19-AUG-1997; 97US-0056730P.
PR	29-JUL-1998; 98WO-US015949.
PR	26-JAN-1999; 99US-00236557.
PR	05-JUN-2001; 2001US-0295558P.
PA	(HUMA-) HUMAN GENOME SCI INC.
XX	
PI	Ruben SM, Feng P, Lafleur DW, Moore PA, Shi Y, Kyaw H, Li Y;
PI	Zeng Z, Carter KC, Endress GA, Wei Y, Fan P, Rosen CA;
XX	
DR	WPI; 2003-874923/81.
XX	
PT	Nucleic acids encoding 83 secreted polypeptides, useful for preventing,
PT	diagnosing and treating disorders related to their aberrant expression
XX	and activity.
XX	
XX	Disclosure; SEQ ID NO 320; 308pp; English.
XX	
CC	The invention relates to an isolated nucleic acid molecule encoding a
CC	secreted protein that is at least 95% identical to a polynucleotide
CC	fragment of any of the nucleotide sequences listed in table 1A of the
CC	specification, which is hybridisable to the nucleotide sequences, a
CC	polynucleotide encoding a polypeptide (or a polypeptide fragment, domain
CC	or epitope of any of the amino acid sequences) listed in table 1A of the
CC	specification, a polynucleotide which is an (allelic) variant of the
CC	nucleotide sequences listed in the specification, a polynucleotide which
CC	encodes a species homologue of the above amino acid sequences, a
CC	polynucleotide capable of hybridising under stringent conditions to any
CC	of the above polynucleotides, where the polynucleotide does not hybridise
CC	under stringent conditions to a nucleic acid molecule having a nucleotide
CC	sequence of only A or T residues. Also included are a recombinant vector
CC	comprising the above nucleic acid molecule, making a recombinant host
CC	cell comprising the above nucleic acid molecule, an isolated polypeptide
CC	comprising a sequence that is at least 95% identical to the polypeptide
CC	(or its fragment, domain, epitope, secreted form, (allelic) variant or
CC	homologue) encoded by the above nucleic acid molecule, an isolated

CC antibody that binds specifically to the above polypeptide, a recombinant
 CC host cell produced by the above method and that expresses the above
 CC polypeptide, making an isolated polypeptide, preventing, treating or
 CC ameliorating a medical condition, diagnosing a pathological condition or
 CC a susceptibility to a pathological condition in a subject, identifying a
 CC binding partner to the above polypeptide, the gene corresponding to the
 CC cDNA sequence given in the specification, and identifying an activity in
 CC a biological assay. The nucleic acid molecule and polypeptide are useful
 CC in diagnosing, preventing, prognosing or treating diseases or disorders
 CC associated with aberrant expression and/or activity of the above
 CC polypeptide, such as neural disorders, immune system disorders, muscular
 CC disorders, reproductive disorders, gastrointestinal disorders, pulmonary
 CC disorders, cardiovascular disorders, renal disorders, proliferative
 CC disorders and/or cancers. In particular, these diseases are systemic
 CC lupus erythematosus, rheumatoid arthritis, multiple sclerosis,
 CC thyroditis, anaemia, Grave's disease, diabetes, hepatitis, asthma,
 CC allergies, nephritis, Parkinson's disease, Alzheimer's disease,
 CC atherosclerosis, myocardial infarction, AIDS and infections. The methods
 CC may be used for identifying agonists and antagonists of the
 CC polynucleotide and polypeptide. The present sequence is a protein from
 CC one of the 83 disclosed secreted protein genes.

Sequence 40 AA;

Query Match 100.0%; Score 15; DB 7; Length 40;
 Best Local Similarity 60.0%; Pred. No. 1.2e+03;
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 RXLXF 8
 DB 2 RLSLF 6

RESULT 42

AB057531
 ID AB057531 standard; protein; 40 AA.

AC AB057531;

DT 29-JUL-2004 (first entry)

DE Human genome derived single exon protein #3765.

KW Human; gene expression; single exon probe; microarray;

KW alternative splicing event; genomic alteration.

OS Homo sapiens.

PN US2003194704-A1.

PD 16-OCT-2003.

PF 03-APR-2002; 2002US-00029386.

PR 03-APR-2002; 2002US-00029386.

PA (PENN/) PENN S G.

PA (RANK/) RANK D R.

PA (HANZ/) HANZEL D K.

PI Penn SG, Rank DR, Hanzel DK;

DR WPI, 2004-119264/12.

PT New human genome-derived single exon nucleic acid probes useful for human

PT gene expression analysis, for identifying or characterizing alternative

PT splicing events, for assessing genomic alterations or as tools for

PT surveying tissues.

PS Claim 45; SEQ ID NO 31165; 80pp; English.

XX The invention relates to a nucleic acid probe for measuring human gene

XX expression, comprising any of the 27,400 fully defined nucleotide

CC sequences in the specification, or their complements or fragments, and
 CC encoding at least 8 amino acids of any of the 688 amino acid sequences
 CC fully defined in the specification. The probe is a single exon probe that
 CC hybridises under high stringency conditions to a nucleic acid molecule
 CC expressed in human cells or tissues. Also included are a spatially-
 CC addressable set of single exon nucleic acid probes for measuring human
 CC gene expression (comprising a plurality of single exon nucleic acid
 CC probes cited above, where each of the plurality of probes is separately
 CC and addressably isolatable or amplifiable from the plurality), a single
 CC exon microarray for measuring human gene expression, a method of
 CC measuring human gene expression, a vector comprising the single exon
 CC probe cited above, an ORF-encoded peptide comprising at least 8
 CC contiguous amino acids of any of the above-mentioned amino acid
 CC sequences (optionally with conservative amino acid substitutions), an
 CC isolated antibody that binds specifically to a peptide cited above, a
 CC method of seiling and/or licensing single exon probes or microarrays to
 CC a customer desiring to measure gene expression, a method of providing
 CC human gene expression data by subscription, and a computer-readable
 CC storage medium which contains a database having a plurality of records
 CC (each record including data on the expression of a single exon probe
 CC cited above. The probe, methods and apparatus are useful in gene
 CC expression analysis. The probes may be used as tools for surveying
 CC tissues to detect the presence of expressed messages that contain their
 CC specific exon, or in constructing genome-derived single exon microarrays.
 CC In addition, the probes are used in identifying and characterising
 CC alternative splicing events, in detecting and characterising gross
 CC alterations in the genomic locus that includes their exon, in assessing
 CC smaller genomic alterations, in printing the synthesis of nucleic acids,
 CC or in expressing the ORF-encoded peptide. The present sequence is a human,
 CC single exon probe protein of the invention. Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from USPTO at
 CC seqdata.uspto.gov/sequence.html?docid=20030194704

Sequence 40 AA;

Query Match 100.0%; Score 15; DB 8; Length 40;
 Best Local Similarity 60.0%; Pred. No. 1.2e+03;
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 RXLXF 8
 DB 22 RLSLF 26

RESULT 43

ADY25664
 ID ADY25664 standard; peptide; 40 AA.

AC ADY25664;

DT 05-MAY-2005 (first entry)

DE Novel human secreted protein-related peptide SeqID320.

KW cancer; cytostatic; arthritis; antiarthritic; asthma; antiasthmatic;

KW acquired immune deficiency syndrome; rheumatoid arthritis; antirheumatic;

KW inflammatory bowel disease; antiinflammatory; gastrointestinal-gen.;

KW sepsis; antibacterial; immunosuppressive; acne; antieborrheic;

KW dermatological; psoriasis; antipsoriatic; atherosclerosis;

KW antiatherosclerotic; cerebrovascular ischemia; cerebroprotective;

KW vasotropic; thrombosis; wound healing; vulnery; Alzheimer's disease;

KW neuroprotective; nootropic; parkinson's disease; antiparkinsonian; autism;

KW obsessive-compulsive disorder; tranquilizer; graft versus host disease;

KW immune disorder; hematological disease; inflammation; infection;

KW hyperproliferative disorders; renal disease; nephropathic disorder;

KW cardiovascular disease; cardiovascular-gen.; respiratory disorder;

KW neurological disease; neuroprotective; endocrine disease;

KW reproductive disorders (general); gynecological.

OS Homo sapiens.

XX US2005037467-A1.

```

XX 17-FEB-2005.
PD
XX 09-SEP-2004; 2004US-00936773.
PF
XX 30-JUL-1997; 97US-0054209P.
XX 30-JUL-1997; 97US-0054211P.
XX 30-JUL-1997; 97US-0054212P.
XX 30-JUL-1997; 97US-0054213P.
XX 30-JUL-1997; 97US-0054214P.
XX 30-JUL-1997; 97US-0054215P.
XX 30-JUL-1997; 97US-0054216P.
XX 30-JUL-1997; 97US-0054217P.
XX 30-JUL-1997; 97US-0054218P.
XX 30-JUL-1997; 97US-0054233P.
XX 30-JUL-1997; 97US-0054236P.
XX 18-AUG-1997; 97US-0055972P.
XX 18-AUG-1997; 97US-0055972P.
XX 18-AUG-1997; 97US-0055972P.
XX 19-AUG-1997; 97US-0056534P.
XX 19-AUG-1997; 97US-0056534P.
XX 19-AUG-1997; 97US-0056543P.
XX 19-AUG-1997; 97US-0056544P.
XX 19-AUG-1997; 97US-0056561P.
XX 19-AUG-1997; 97US-0056727P.
XX 19-AUG-1997; 97US-0056730P.
XX 29-JUL-1998; 98WO-US015949.
XX 26-JAN-1999; 99US-00236557.
XX 21-SEP-2000; 2000US-00666984.
XX 30-MAR-2001; 2001US-00820649.
XX 05-JUN-2001; 2001US-0295558P.
XX 04-JUN-2002; 2002US-00160162.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX Ruben SM, Feng P, Lafleur DW, Moore PA, Shi Y, Kyaw H, Li Y;
PI Zeng Z, Carter KC, Endress GA, Wei Y, Fan P, Rosen CA;
PI WPI; 2005-161941/17.
XX
XX New nucleic acid molecule encoding human secreted protein, useful for
PT preventing, treating, or ameliorating immune system, blood, inflammatory,
PT infectious, cardiovascular, respiratory, neurological, endocrine, or
PT reproductive disorders.
XX
XX Disclosure; SEQ ID NO 320; 309pp; English.
XX
XX The invention comprises the amino acid and coding sequences of novel
CC human secreted proteins and peptides. The DNA and protein sequences of
CC the invention are useful for treating or ameliorating: cancer, arthritis,
CC asthma, AIDS, rheumatoid arthritis, inflammatory bowel disease, sepsis,
CC acne, psoriasis, atherosclerosis, stroke, thrombosis, wound healing,
CC Alzheimer's Disease, Parkinson's Disease, autism, obsessive compulsive
CC disorder, graft-versus-host diseases, autism, obsessive compulsive
CC disorders, inflammatory conditions, immune system disorders, blood
CC hyperproliferative disorders, renal disorders, cardiovascular disorders,
CC respiratory disorders, neurological disorders, endocrine disorders, or
CC reproductive disorders. The present sequence is that of a human peptide
CC which is related to the novel human secreted proteins/peptides of the
CC invention.
XX
XX Sequence 40 AA;
SQ
Query Match 100.0%; Score 15; DB 9; Length 40;
Best Local Similarity 60.0%; Pred. No. 1.2e+03;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 4 RXLXF 8
| | |
Db 2 RSLSF 6

```

```

ID AAY25813 standard; protein; 43 AA.
XX
XX AAY25813;
AC
XX 04-OCT-1999 (first entry)
DT
XX Human secreted protein fragment encoded from gene 23.
DE
XX Secreted protein; human; treatment; diagnosis; therapy; cancer; tumour;
XX neurodegenerative disorder; developmental abnormality; blood disorder;
XX fetal deficiency; blood disorder; leukemia; immune system; inflammation;
XX autoimmune disease; hepatic disease; renal disease; allergy; restenosis;
XX ischaemic shock; Alzheimer's disease; cognitive disorder; schizophrenia;
XX cardiovascular disorder; wound healing; stroke; arthritis; obesity;
XX asthma; sepsis; acne; psoriasis; transplant rejection; infection; AIDS;
XX metabolic disorder.
XX
XX Homo sapiens.
OS
XX
XX W09938881-A1.
XX
XX 05-AUG-1999.
XX
XX 27-JAN-1999; 99WO-US001621.
XX
XX 30-JAN-1998; 98US-0073159P.
XX 30-JAN-1998; 98US-0073160P.
XX 30-JAN-1998; 98US-0073161P.
XX 30-JAN-1998; 98US-0073162P.
XX 30-JAN-1998; 98US-0073164P.
XX 30-JAN-1998; 98US-0073165P.
XX 30-JAN-1998; 98US-0073167P.
XX 30-JAN-1998; 98US-0073170P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX Ruben SM, Ferrite AM, Rosen CA, Florence KA, Carter KC, Soppet DR;
PI Yu G, Florence C, Young P, Ni J, Feng P, Endress GA, Janat F;
PI WPI; 1999-469315/39.
XX
XX N-PSDB; AAZ00432.
XX
XX New isolated human genes and the secreted polypeptides they encode useful
PT in, e.g. treatment of Alzheimer's.
PT
XX Disclosure; Page 355; 393pp; English.
XX
XX This invention describes novel human genes (see AAZ00410-200477) and the
CC secreted proteins (see AAY25711-Y25778) and fragments (see AAY25779-
CC Y25907) they encode. The polynucleotides and their corresponding secreted
CC polypeptides are useful for preventing, treating or ameliorating medical
CC conditions e.g. by protein or gene therapy. Also pathological conditions
CC can be diagnosed by determining the amount of the new polypeptides in a
CC sample or by determining the presence of mutations in the new
CC polynucleotides. Specific uses are described for each of the 67
CC polynucleotides of the invention, based on which tissues they are most
CC highly expressed in, and include developing products for the diagnosis or
CC treatment of cancer, tumours, neurodegenerative disorders, developmental
CC abnormalities and fetal deficiencies, blood disorders, leukemias,
CC diseases of the immune system, autoimmune diseases, hepatic and renal
CC disease, inflammation, allergies, ischaemic shock, Alzheimer's and
CC cognitive disorders, schizophrenia, restenosis, cardiovascular disorders,
CC wound healing, stroke, arthritis, obesity, asthma, sepsis, acne,
CC psoriasis, transplant rejection, metabolic disorders, infections and
CC AIDS. The polypeptides are also useful for identifying their binding
CC partners.
XX
XX Sequence 43 AA;
SQ
Query Match 100.0%; Score 15; DB 2; Length 43;
Best Local Similarity 60.0%; Pred. No. 1.3e+03;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 4 RXLXF 8
Db 28 RTLTF 32

RESULT 45
AAU86594
ID AAU86594 standard; protein; 45 AA.
XX
AC AAU86594;
XX
DT 21-MAY-2002 (first entry)
XX
DE Novel human connective tissue related polypeptide #160.
XX
KW Human; connective tissue related disorder; cancer; cyrostatic.
XX
OS Homo sapiens.
XX
PN WO200155343-A1.
XX
FD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001322.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-020515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-022513P.
PR 14-AUG-2000; 2000US-022514P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-022547P.
PR 14-AUG-2000; 2000US-022557P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-022679P.
PR 22-AUG-2000; 2000US-022681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 23-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.

PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.

PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251888P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PI Rosen CA, Barash SC, Ruben SM;
XX WPI, 2001-565190/63.
DR N-PSDB; ABR41772.
XX
XX Nucleic acid encoding novel connective tissue associated polypeptides,
PT used in diagnosing, preventing, treating or ameliorating a disorder such
PT as cancer or rheumatoid arthritis.
XX
XX Claim 11; SEQ ID NO 659; 673pp; English.
XX
XX The present invention relates to the isolation of novel human connective
CC tissue related polypeptides and the polynucleotide (cDNA and genomic)
CC sequences encoding them. The sequences of the invention are useful in the
CC diagnosis, treatment, prevention and/or prognosis of diseases associated
CC with connective tissue(s), including cancer. The polynucleotide sequences
CC of the invention are also useful in gene therapy. A086435-A086923
CC represent the novel human connective tissue related polypeptides. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 45 AA;
SQ
Query Match 100.0%; Score 15; DB 4; Length 45;
Best Local Similarity 60.0%; Pred. No. 1.3e+03;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 4 RXLXF 8
Db 27 RALSF 31
RESULT 46
ADBS9928
ID ADBS9928 standard; protein; 45 AA.
XX
XX ADBS9928;
AC
XX 04-DEC-2003 (first entry)
DT
XX
DE Connective tissue antigen (CTA) #160.
XX
XX cytosolic; neuroprotective; nootropic; antiparkinsonian; cardiovascular;
KW antiarteriosclerotic; immunosuppressive; antirheumatic; antichronic;
KW antiinflammatory; antiallergic; antisthmatic; dermatological;
KW nephrotropic; virocidic; fungicide; antibacterial; antiparasitic;
KW gene therapy; ds; connective tissues disorder; rheumatoid arthritis;
KW systemic lupus erythematosus; scleroderma; Sjogren's syndrome; cancer;
KW cancer metastasis; neoplasia; leukemia; neurodegenerative disorder;
KW Alzheimer's disease; Parkinson's disease; cardiovascular disease;
KW atherosclerosis; myocarditis; cardiopulmonary bypass complication;
KW autoimmune disease; multiple sclerosis; allergic reaction; asthma;
KW rhinitis; eczema; inflammatory condition; Crohn's disease; nephritis;
KW gastrointestinal disorder; inflammatory bowel disease;

KW organ transplant rejection; immune system disorder; Bruton's disease;
KW X-linked lymphoproliferative syndrome;
KW B-cell lymphoproliferative disorder; HIV; AIDS; infection;
KW chromosome identification; chromosome mapping; connective tissue antigen;
KW CTA.
XX
XX Homo sapiens.
XX OS
XX US2003054375-A1.
XX
XX 20-MAR-2003.
XX
XX 07-MAR-2002; 2002US-00092154.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 24-FEB-2000; 2000US-0184664P.
XX 02-MAR-2000; 2000US-0186350P.
XX 16-MAR-2000; 2000US-0189874P.
XX 17-MAR-2000; 2000US-0190076P.
XX 18-APR-2000; 2000US-0198123P.
XX 19-MAY-2000; 2000US-0205515P.
XX 07-JUN-2000; 2000US-0209467P.
XX 28-JUN-2000; 2000US-0214886P.
XX 30-JUN-2000; 2000US-0215135P.
XX 07-JUL-2000; 2000US-0216647P.
XX 07-JUL-2000; 2000US-0216880P.
XX 11-JUL-2000; 2000US-0217487P.
XX 11-JUL-2000; 2000US-0217496P.
XX 14-JUL-2000; 2000US-0218290P.
XX 26-JUL-2000; 2000US-0220963P.
XX 26-JUL-2000; 2000US-0220964P.
XX 14-AUG-2000; 2000US-0224518P.
XX 14-AUG-2000; 2000US-0224519P.
XX 14-AUG-2000; 2000US-0225213P.
XX 14-AUG-2000; 2000US-0225214P.
XX 14-AUG-2000; 2000US-0225266P.
XX 14-AUG-2000; 2000US-0225267P.
XX 14-AUG-2000; 2000US-0225268P.
XX 14-AUG-2000; 2000US-0225270P.
XX 14-AUG-2000; 2000US-0225447P.
XX 14-AUG-2000; 2000US-0225757P.
XX 14-AUG-2000; 2000US-0225758P.
XX 14-AUG-2000; 2000US-0225759P.
XX 18-AUG-2000; 2000US-0226279P.
XX 22-AUG-2000; 2000US-0226681P.
XX 22-AUG-2000; 2000US-0226682P.
XX 22-AUG-2000; 2000US-0227182P.
XX 23-AUG-2000; 2000US-0227009P.
XX 30-AUG-2000; 2000US-0228894P.
XX 01-SEP-2000; 2000US-0229287P.
XX 01-SEP-2000; 2000US-0229343P.
XX 01-SEP-2000; 2000US-0229344P.
XX 01-SEP-2000; 2000US-0229345P.
XX 05-SEP-2000; 2000US-0229509P.
XX 05-SEP-2000; 2000US-0229513P.
XX 06-SEP-2000; 2000US-0230437P.
XX 06-SEP-2000; 2000US-0230438P.
XX 06-SEP-2000; 2000US-0231242P.
XX 08-SEP-2000; 2000US-0231243P.
XX 08-SEP-2000; 2000US-0231244P.
XX 08-SEP-2000; 2000US-0231413P.
XX 08-SEP-2000; 2000US-0231414P.
XX 08-SEP-2000; 2000US-0232080P.
XX 08-SEP-2000; 2000US-0232081P.
XX 12-SEP-2000; 2000US-0231968P.
XX 14-SEP-2000; 2000US-0232397P.
XX 14-SEP-2000; 2000US-0232398P.
XX 14-SEP-2000; 2000US-0232399P.
XX 14-SEP-2000; 2000US-0232400P.
XX 14-SEP-2000; 2000US-0232401P.
XX 14-SEP-2000; 2000US-0233063P.
XX 14-SEP-2000; 2000US-0233064P.

PR 14-SEP-2000; 2000US-0231065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239359P.
PR 13-OCT-2000; 2000US-0239377P.
PR 13-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251566P.
PR 08-DEC-2000; 2000US-0251568P.
PR 08-DEC-2000; 2000US-0251569P.

PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PR 17-JAN-2001; 2001US-00764847.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM, Barash SC;
XX WPI; 2003-634869/60.
XX N-PSDB; ADB59439.
XX
XX New connective tissue-related polypeptides and polynucleotides, useful
XX for treating, preventing and/or prognosing e.g. disorders of connective
XX tissue, (e.g. rheumatoid arthritis), cancers, cancer metastases and/or
XX neoplasias.
XX
XX Claim 11; SEQ ID NO 659; 248bp; English.
XX
XX The invention describes an isolated nucleic acid molecule (I), which
XX comprises a sequence that is at least 95 % identical to a connective
XX tissue-related polynucleotide encoding connective tissue antigens (CTA).
XX The polypeptide or polynucleotide is useful for preventing, treating, or
XX ameliorating medical conditions in a mammal. The connective tissue
XX polypeptides, polynucleotides and antibodies are particularly useful for
XX treating, preventing and/or prognosing disorders of connective tissues
XX (e.g. rheumatoid arthritis, discoid and systemic lupus erythematosus,
XX scleroderma, or Sjogren's syndrome), cancers, cancer metastases and/or
XX neoplasias (e.g. leukemias), neurodegenerative disorders (e.g.
XX Alzheimer's disease, or Parkinson's disease), cardiovascular diseases
XX (e.g. atherosclerosis, myocarditis or cardiopulmonary bypass
XX complications), autoimmune diseases (e.g. systemic lupus erythematosus,
XX rheumatoid arthritis, or multiple sclerosis), allergic reactions (e.g.
XX
XX
Query Match 100.0%; Score 15; DB 7; Length 45;
Best Local Similarity 60.0%; Pred. No. 1.3e+03;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 4 RFLXF 8
DB 27 RALSF 31
RESULT 47
ADN42941
ID ADN42941 standard; protein; 45 AA.
XX
XX ADN42941;
XX
XX 29-JUL-2004 (first entry)
XX
XX Human secreted protein SECP-2.
XX
XX
XX Cytostatic; anorectic; immunosuppressive; gene therapy; SECP-antagonist;
XX SECP-agonist; secreted protein; SECP; autoimmune disorder; obesity;
XX cancer; human; SECP-2.
XX
XX Homo sapiens.
XX
XX WO2004037987-A2.
XX
XX 06-MAY-2004.
XX
XX 22-OCT-2003; 2003WO-US033491.
XX
XX 22-OCT-2002; 2002US-0420720P.
XX 07-NOV-2002; 2002US-0425207P.
XX 15-NOV-2002; 2002US-042679P.
XX 19-NOV-2002; 2002US-0427871P.
XX 06-JAN-2003; 2003US-0438551P.
XX 17-JAN-2003; 2003US-0441144P.
XX 27-JAN-2003; 2003US-0443135P.

XX (INCY-) INCYTE CORP.
PA
XX
XX Baughn MR, Becha SD, Bhatia UG, Blake JJ, Burrill JD, Chawla NK;
PI Chien D, Elliott VS, Emerling EM, Faveiro KD, Hafalia AJA;
PI Harmsen BW, Ison CH, Khare R, Lee SY, Lu DM;
PI Marquis JP, Murage J, Nguyen DB, Rankumar J, Richardson TW;
PI Swarnakar A, Tang TY, Tran UK, Wang JT, Yue H, Zheng W;
XX WPI: 2004-365505/34.
DR N-PSDB; ADN43016.
XX
XX New human secreted protein (SECP) polypeptide, useful for preparing a
PT composition for treating a disease associated with decreased expression
PT or overexpression of functional SECP e.g., autoimmune disorders, obesity
PT or cancer.
XX
XX Claim 1; SEQ ID NO 2; 289bp; English.
XX
XX The invention describes an isolated human secreted protein (SECP)
CC polypeptide. Also described are: an isolated polynucleotide encoding the
CC polypeptide; a recombinant polynucleotide comprising a promoter sequence
CC operably linked to the polynucleotide; a cell transformed with the
CC recombinant polynucleotide; a transgenic organism comprising the
CC recombinant polynucleotide; a method of producing the polypeptide; an
CC isolated antibody that specifically binds to the polypeptide; a method of
CC detecting a target polynucleotide in a sample; a method for treating a
CC disease or condition associated with decreased expression or
CC overexpression of functional SECP; a method of screening a compound for
CC effectiveness as an agonist or antagonist of the polypeptide or in
CC altering expression of the target polynucleotide; a method of screening a
CC compound that specifically binds to, or that modulates the activity of,
CC the polypeptide; a method of assessing toxicity of a test compound; a
CC diagnostic test for a condition or disease associated with the expression
CC of SECP in a biological sample; a method of diagnosing a condition or
CC disease associated with the expression of SECP in a subject; a
CC composition comprising the antibody and a carrier, or the polypeptide, or
CC agonist or antagonist compound and an excipient; a method of preparing a
CC polyclonal or monoclonal antibody; a method of detecting the polypeptide
CC in a sample; a method of purifying the polypeptide; a method of
CC generating an expression profile of a sample that contains
CC polynucleotides; and an array comprising different nucleotide molecules
CC affixed in distinct physical locations on a solid substrate, where at
CC least one of the nucleotide molecules comprises a first oligonucleotide
CC or polynucleotide sequence specifically hybridizable with at least 30
CC contiguous nucleotides of the target polynucleotide. The polypeptide is
CC useful for preparing a composition for diagnosing or treating a disease
CC or condition associated with decreased expression or overexpression of
CC functional SECP e.g. autoimmune disorders, obesity or cancer. This is the
CC amino acid sequence of a human secreted protein.
XX
XX Sequence 45 AA:
SQ
Query Match 100.0%; Score 15; DB 8; Length 45;
Best Local Similarity 60.0%; Pred. No. 1.3e+03;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 4 RXLXF 8
Db 40 RALSF 44

RESULT 48
ABG26048
ID ABG26048 standard; protein; 49 AA.
XX
XX ABG26048;
AC
XX
XX 18-FEB-2002 (first entry)
DT
XX
XX Novel human diagnostic protein #26039.
DE
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX Homo sapiens.
XX
XX WO200175067-A2.
XX
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US008631.
XX
XX 31-MAR-2000; 2000US-00540217.
XX
XX 23-AUG-2000; 2000US-00649167.
XX
XX (HXSE-) HXSE INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX
XX WPI: 2001-639362/73.
XX
XX N-PSDB; AAS90235.
XX
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
XX Claim 20; SEQ ID NO 56407; 103bp; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridization probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polynucleotide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 49 AA:
SQ
Query Match 100.0%; Score 15; DB 4; Length 49;
Best Local Similarity 60.0%; Pred. No. 1.4e+03;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 4 RXLXF 8
Db 17 RALSF 21

RESULT 49
ADS98736
ID ADS98736 standard; protein; 49 AA.
XX
XX ADS98736;
AC
XX
XX 30-DEC-2004 (first entry)
DT
XX
XX Protein factor discovery related human contig polypeptide, SEQ ID 1000.
DE
XX anti-inflammatory; cytostatic; antimicrobial; gene therapy; inflammation;
XX leukemia; nervous system disorder; infection.
KW

OS	Homo sapiens.
XX	
FN	WO2004087874-A2.
XX	
PD	14-OCT-2004.
XX	
PF	24-MAR-2004; 2004WO-US009202.
XX	
PR	28-MAR-2003; 2003US-045882A.P.
XX	
PA	(NUVE-) NUVELO INC.
XX	
PA	(DRMA//) DRMANNA R.T.
PI	Tang YF, Zhou P, Wang J, Wang ZW, Hu T;
XX	
DR	WPI: 2004-737686/72.
DR	N-PSDB; ADS98396.
XX	
PT	New polynucleotides encoding a polypeptide with biological activity,
PT	useful for treating inflammation, leukemias, nervous system disorders, or
PT	infections.
XX	
PS	Example 3; SEQ ID NO 1000; 253bp, English.
XX	
CC	The invention relates to a novel isolated polynucleotide comprising any
CC	of the 235 nucleotide sequences described in the specification. The
CC	invention further comprises: an isolated polynucleotide encoding a
CC	polypeptide with biological activity, where the polynucleotide hybridizes
CC	to one of the 235 novel polynucleotides under stringent hybridization
CC	conditions, or having greater than about 9% sequence identity with the
CC	novel polynucleotide; a vector comprising a novel polynucleotide; an
CC	expression vector comprising the novel polynucleotide; a host cell
CC	genetically engineered to comprise the novel polynucleotide, which can be
CC	operatively associated with a regulatory sequence that modulates
CC	expression of the polynucleotide in the host cell; an isolated
CC	polypeptide encoded by the novel polynucleotide, or a polynucleotide
CC	hybridizing under stringent conditions to the novel polynucleotide; a
CC	composition comprising the polypeptide and a carrier; an antibody
CC	directed against the polypeptide; a method for detecting the novel
CC	polynucleotide in a sample; a method for detecting the polypeptide in a
CC	sample; a method for identifying a compound that binds to the polypeptide
CC	; a method for producing the polypeptide; an isolated polypeptide
CC	comprising any of the 235 amino acid sequences described in the
CC	specification; and a collection of polynucleotides comprising of at least
CC	one of the polynucleotides cited above. The polypeptides and
CC	polynucleotides of the invention have antiinflammatory, cytostatic, and
CC	antimicrobial activities. The novel polynucleotide may be used to treat
CC	disorders by gene therapy. The polypeptides and polynucleotides are
CC	useful for treating inflammation, leukaemias, nervous system disorders,
CC	or infections. This sequence represents the polypeptide encoded by
CC	contiguous DNA derived from one of the 235 novel isolated polynucleotides
CC	of the invention.
XX	
SQ	Sequence 49 AA:
OY	Query Match 100.0%; Score 15; DB 8; Length 49; Best Local Similarity 60.0%; Pred.No.1.4e+03; Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0. 4 RXLXF 8 17 RALAF 21
RESULT 50	
AAM86216	
ID	AAM86216 standard; protein; 50 AA.
XX	
AC	AAM86216;
XX	
DT	07-NOV-2001 (first entry)
XX	
DE	Human immune/haematopoietic antigen SEQ ID NO:13809.

XX	Human; immune; haematopoietic; immune/haematopoietic antigen; cancer
KM	cytostatic; gene therapy; vaccine; metastasis.
XX	
OS	Homo sapiens.
XX	
PM	WO200157182-A2.
XX	
PD	09-AUG-2001.
XX	
XX	
PF	17-JAN-2001; 2001WO-US001354.
XX	
PR	31-JAN-2000; 2000US-0179065P.
PR	04-FEB-2000; 2000US-0180628P.
PR	24-FEB-2000; 2000US-0184664P.
PR	02-MAR-2000; 2000US-0186350P.
PR	16-MAR-2000; 2000US-0189874P.
PR	17-MAR-2000; 2000US-019076P.
PR	18-APR-2000; 2000US-0198137P.
PR	19-MAY-2000; 2000US-0205515P.
PR	07-JUN-2000; 2000US-0209467P.
PR	26-JUN-2000; 2000US-0214886P.
PR	30-JUN-2000; 2000US-0215135P.
PR	07-JUL-2000; 2000US-0216647P.
PR	07-JUL-2000; 2000US-0216880P.
PR	11-JUL-2000; 2000US-0217487P.
PR	11-JUL-2000; 2000US-0217496P.
PR	14-JUL-2000; 2000US-0218293P.
PR	26-JUL-2000; 2000US-0220963P.
PR	26-JUL-2000; 2000US-0220964P.
PR	14-AUG-2000; 2000US-0224518P.
PR	14-AUG-2000; 2000US-0224519P.
PR	14-AUG-2000; 2000US-0225213P.
PR	14-AUG-2000; 2000US-0225214P.
PR	14-AUG-2000; 2000US-0225267P.
PR	14-AUG-2000; 2000US-0225268P.
PR	14-AUG-2000; 2000US-0225270P.
PR	14-AUG-2000; 2000US-0225447P.
PR	14-AUG-2000; 2000US-0225757P.
PR	14-AUG-2000; 2000US-0225758P.
PR	14-AUG-2000; 2000US-0225759P.
PR	18-AUG-2000; 2000US-0226279P.
PR	22-AUG-2000; 2000US-0226681P.
PR	22-AUG-2000; 2000US-0226868P.
PR	23-AUG-2000; 2000US-0227182P.
PR	23-AUG-2000; 2000US-0227009P.
PR	30-AUG-2000; 2000US-0228924P.
PR	01-SEP-2000; 2000US-0228287P.
PR	01-SEP-2000; 2000US-0229343P.
PR	01-SEP-2000; 2000US-0229344P.
PR	01-SEP-2000; 2000US-0229345P.
PR	05-SEP-2000; 2000US-0229509P.
PR	05-SEP-2000; 2000US-0229513P.
PR	06-SEP-2000; 2000US-0230437P.
PR	06-SEP-2000; 2000US-0230438P.
PR	08-SEP-2000; 2000US-0231242P.
PR	08-SEP-2000; 2000US-0231243P.
PR	08-SEP-2000; 2000US-0231244P.
PR	08-SEP-2000; 2000US-0231414P.
PR	08-SEP-2000; 2000US-0231415P.
PR	08-SEP-2000; 2000US-0232080P.
PR	08-SEP-2000; 2000US-0232081P.
PR	12-SEP-2000; 2000US-0231968P.
PR	14-SEP-2000; 2000US-0232397P.
PR	14-SEP-2000; 2000US-0232398P.
PR	14-SEP-2000; 2000US-0232399P.
PR	14-SEP-2000; 2000US-0232400P.
PR	14-SEP-2000; 2000US-0233063P.
PR	14-SEP-2000; 2000US-0233064P.
PR	14-SEP-2000; 2000US-0233065P.
PR	21-SEP-2000; 2000US-0234223P.

PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 25-SEP-2000; 2000US-0234998P.
 PR 26-SEP-2000; 2000US-0234984P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235836P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239935P.
 PR 13-OCT-2000; 2000US-0239937P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241825P.
 PR 01-NOV-2000; 2000US-024617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249246P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.

PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 XX
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Roese CA, Barash SC, Ruben SM;
 XX WPI, 2001-483426/52.
 DR N-PSDB; AAK58997.
 PT
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 XX useful for preventing, diagnosing and/or treating cancers and metastasis.
 PS
 PS Claim 11; SEQ ID NO 13809; 3071pp + Sequence Listing; English.
 XX
 CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
 CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patients own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting the
 CC nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/haematopoietic-related diseases, especially
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/haematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
 CC represent sequences used in the exemplification of the present invention
 CC
 XX
 SQ Sequence 50 AA:

Query Match 100.0%; Score 15; DB 4; Length 50;
 Best Local Similarity 60.0%; Pred. No. 1.5e+03;
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RXLXF 8
 | | |
 Db 6 RSLAF 10

Search completed: May 4, 2006, 06:05:50
 Job time : 196 secs

THIS PAGE BLANK (USPTO)